THE UNIVERSITY OF
WARWICK

Optical Free Space Feedforward Non-linearity Correction System

By

Cameron Sweet

Submitted for the Degree of Doctor of Philosophy in Engineering

UNVERSITY OF WARWICK
SCHOOL OF ENGINEERING

September 2003

Table of Contents

Table of Contents ii

List of illustrations and tables iv

Abbreviations vi

Acknowledgements viii

Declaration x

Abstract xi

Chapter 1 – Introduction 1

1.1 Background 1

1.2 GSM, UMTS and 3G 3

1.3 Remote Antennas 8

1.4 Optical Free Space and Optical Fibre 12

Chapter 2 – The Non-Linearities present and the techniques available to correct them. 15

2.1 Link design specifications 15

2.2 Noise 18

2.3 Properties of the Air Interface 23

2.4 Non-linear Laser Diode Distortion 29

2.5 Non-linear correction techniques 34

2.6 Feedforward Linearisation for Optical Free Space 41

Chapter 3 – Taylor's Series, Perturbation Analysis and Volterra Series Expansions 44

3.1 The Taylor series 44

3.2 Perturbation analysis 51

3.3 Volterra series analysis 55

3.4 Advanced Volterra series analysis, including carrier transport effects 62
3.5 Simulations at low frequency 68
3.6 Summary and conclusions 69

Chapter 4 – Noise Analysis 71
4.1 Standard noise analysis (with 1:1 coupling ratio) 71
4.2 Advanced noise analysis (with a 9:1 coupling ratio) 80
4.3 Summary and conclusions 82

Chapter 5 – Practical Work 83
5.1 Stage one set-up, with electrical correction path 83
5.2 Stage one results 91
5.3 Stage two set-up, with optical correction path 94
5.4 Stage two results 96
5.5 Noise analysis 100
5.6 Conclusion 100

Chapter 6 – Conclusions and further work 101

References 106

Appendix A – Perturbation analysis 111
Appendix B – Volterra Analysis 115
Appendix C – Matlab Simulation Results 126
Appendix D – Data sheets and circuit diagrams 131

List of illustrations and tables

Chapter 1

Figure 1.1	Page 5	UMTS concept, courtesy of the IEE
Table 1.1	Page 6	General comparison of 2G and 3G parameters
Table 1.2	Page 7	Main FMA parameters
Figure 1.2	Page 9	Cell distribution with frequency reuse
Figure 1.3	Page 10	Principle of remote antennas, over optical fibre and free space

Chapter 2

Figure 2.1	Page 18	Spurious Free Dynamic Range
Figure 2.2	Page 24	Atmospheric transmission of IR Spectrum
Table 2.1	Page 27	Visibility and attenuation for different degrees of haze and fog
Figure 2.3	Page 30	Light verses current characteristic, with the ideal in light blue
Figure 2.4	Page 33	Two tone intermodulation spectrum, with 3^{rd} and 5^{th} order terms
Figure 2.5	Page 35	Optical Feedback technique
Figure 2.6	Page 36	Phase Shift technique
Figure 2.7	Page 37	The Predistortion technique
Figure 2.8	Page 39	The Feedforward technique
Figure 2.9	Page 40	The Quasi-feedforward technique
Figure 2.10	Page 42	Optical Free Space Feedforward technique

Chapter 3

Figure 3.1	Page 48	Results for Taylor's series optical feedforward non-linearity correction simulation
Figure 3.2	Page 51	Expanded frequency scale Taylor's series simulation results

Figure 3.3	Page 54	Results for perturbation analysis of the optical feedforward non-linearity correction simulation
Figure 3.4	Page 56	Representation of a weak non-linear system characterised by a Volterra series
Figure 3.5	Page 61	Results for Volterra series analysis of the optical feedforward non-linearity correction simulation
Figure 3.6	Page 64	Schematic of the three rate equation model of carrier transport in QW laser, with the SCH and barrier region in turquoise and the QW region in purple
Figure 3.7	Page 67	Results for Volterra series analysis of the optical feedforward non-linearity correction simulation, including carrier transport effects
Figure 3.8	Page 69	Results of the Taylor's series and Volterra series simulation at modulation frequencies of 20MHz and 21MHz

Chapter 4

Figure 4.1	Page 72	Unit bandwidth noise source schematic of the feedforward correction system of figure 2.10
Figure 4.2	Page 76/77	System noise outputs against variation of I_1
Figure 4.3	Page 78	System noise outputs against variation of I_2
Figure 4.4	Page 79	System noise outputs against variation of RIN_1
Figure 4.5	Page 81	System noise outputs against inputs of I_2 and RIN_1 for a coupling ratio of 0.1

Chapter 5

Figure 5.1	Page 84	First stage of practical system realisation, with electrical correction path
Figure 5.2	Page 87	Illustration of a pellicle beam splitter compared to conventional beam splitter
Figure 5.3	Page 88	DTIRC 3-D diagram (courtesy of Roberto Remirez-Inguez, University of Warwick)

Figure 5.4	Page 91	Power supply and modulation supply unit for the laser diodes and op-amps
Figure 5.5	Page 92/93	Results of the stage 1, (electrical error signal) system, with; a) pre-corrected signal, b) error signal and c) corrected signal
Figure 5.6	Page 94	Second stage of practical system realisation, with optical correction path
Figure 5.7	Page 95	Stage one optical components
Figure 5.8	Page 96	Final system set-up
Figure 5.9	Page 97/98	Results of the stage 2, (optical error signal) system, with; a) pre-corrected signal and b) corrected signal
Figure 5.10	Page 98	Spectrum of the error signal
Figure 5.11	Page 99	Spectrum of the pre-corrected optical signal
Figure 5.12	Page 99	Spectrum of the corrected optical signal

Chapter 6

Figure 6.1	Page 102	Example of a commercial roof-top system, with signal blocking panel

Abbreviations

2G – Second Generation

3G – Third Generation

AC – Alternating Current

ACTS – Advanced Communications Technologies and Services

AM – Amplitude Modulation

BER – Bit Error Rate

BOQAM – Binary Offset Quadrature Amplitude Modulation

BTS – Base Transceiver Station

CATV – Cable Access Television

CDMA – Code Division Multiple Access

CNR – Carrier to Noise Ratio

CPC – Compound Parabolic Concentrators

CSO – Composite Second Order

CTB – Composite Triple Beat

DAS – Distributed Antenna System

DC – Direct Current

DL – Down Link

DR – Dynamic Range

DTIRC – Dielectric Totally Internally Reflecting Concentrator

EDGE – Enhanced Data rates for GSM Evolution

EIN – Equivalent Input Noise

ETSI – European Telecommunications Standards Institute

FM – Frequency Modulation

FMA – FRAMES multiple Access

FP – Fabry Perot

FRAMES – Future Radio Multiple Access System

GMSK – Gaussian Minimum Shift Keying

GPRS – General Packet Radio Service

GSM – Global System for Mobile

HSCSD – High Speed Circuit Switched Data

IMD – Inter Modulation Distortion

IP – Internet Protocol

IR – Infra Red

ITU – International Telecommunications Union

LD – Laser Diode

LSHB – Longitudinal Spatial Hole Burning

MQW – Multi Quantum Well

NEB – Noise Equivalent Bandwidth

NF – Noise Figure

OMI – Optical Modulation Index

PCB – Printed Circuit Board

PD – Photo Diode

PSHC – Phase-Shift Harmonic Cancellation

QAM – Quadrature Amplitude Modulation

QOQAM – Quaternary Offset Quadrature Amplitude Modulation

QPSK – Quaternary Phase Shift Keying

QW – Quantum well

RF – Radio Frequency

RIN – Relative Intensity Noise

SCH – Separate Confinement Heterostructure

SCM – Sub Carrier Multiplex

SDH – Synchronous Digital Hierarchy

SFDR – Spurious Free Dynamic Range

TDD – Time Division Duplex

TD-CDMA – Time Division - Code Division Multiple Access

TDMA – Time Division Multiple Access

UL – Up Link

UMTS – Universal Mobile Telecommunications System

W-CDMA – Wide band - Code Division Multiple Access

Acknowledgements

First and foremost I would like to thank my supervisor, Prof. Roger J. Green for his guidance, patience and encouragement throughout this work, without which, it would not have got far.

I wish to express my gratitude to the EPSRC for their funding that has allowed me to undertake this study.

Within the School of Engineering, I would like to thank the following. My colleagues in the Communications and Signal Processing Research Group for their camaraderie. The motley crew of the OEL, especially Keith and Paul, for their warm welcome when I first arrived at Warwick and their continuing friendship and advice since. All the academics and technicians I have crossed paths with and especially Dave and Ian for their patience during the practical stage of the project. Finally I would like to thank Clive for his many fine chats about rugby and especially Anne for helping keep me sane.

I wish to thank all my friends that I shared the mighty 21 Kenilworth Rd with, again for the endless fun, friendship and support when the going got tough, especially Joe and George.

I wish to extend my eternal thanks to all my friends from my pre-Warwick life, again for their timeless friendship and support, especially Pob, Tim and Col who helped me believe enough in myself to take on this madness in the first place, thanks chaps.

I would like to thank the late Mr Larry Trout, my mentor at Marconi Communications who encouraged me to get out of the comfort zone and push myself, which results in this thesis

The penultimate thanks go to all my family with extra mentions for my brother Anton, my Dad Nick and especially my Mum Jo, who have backed me all the way during these last few years, both morally and financially.

My last thanks goes to a remarkable young lady who has had to put up with all the highs and lows of the last four and a half years. She never let me get down about my research, encouraging me all the way and looking after me when I was ill. I could not of completed this work without her, therefore, I dedicate this work to her, to my lady Lisa.

Declaration

I hereby declare that the work contained within this thesis is the author's own. Where collaborative work has taken place it will be stated and collaborates listed. This thesis has not been submitted for a degree at any other university.

Work from this thesis has been presented at PREP 2001, 9th to 11th April 2001, Keele, England, PREP 2002, 17th to 19th April 2002, Nottingham, England, the IEEE Wireless Design Conference, 16th to 17th May 2002, London and the IEE 1st International conference on Photonic Access Technologies, 17th to 18th December 2002, London.

Abstract

Recent years have seen unprecedented growth in the popularity and deployment of mobile phones. As this continues, so the strain on existing mobile cellular radio network has also increased, leading to the need to investigate new technologies to relieve this pressure. The problem is being further exacerbated by the introduction of the 3rd generation of mobile communications, otherwise known as UMTS (Universal Mobile Telecommunications System), with the aim of offering multimedia services on pocket sized portable receivers. A major cost of the mobile radio network, in terms of both financial and social/environmental aspects, is the need apparent need for more base transceiver stations (BTS), due to the increased number of services, and the density of them.

Therefore, judicious use of fewer, but more "intelligent" base stations, thereby reducing the overall system costs, and extra flexibility in the design of mobile cells would be desirable. This can be achieved by having the BTS antennas remotely positioned from the BTS by transmitting the radio signals down an optical fibre or, as in this project, over free space. The main application for this is in densely urban heavy use areas, where there is extensive reuse of both cell and cell cluster. This, along with building shadowing, would require a BTS on every corner, and where extra cell design flexibility would be desirable. Also, in remote rural areas, where various natural features, such as rivers or mountains can cause similar cell design problems, there is a need for this flexibility.

The problem with this requirement is that the electrical to optical conversion process, involving a laser diode driver unit, is inherently non-linear, and, unless this is resolved, the desired signal will become unusable due to distortion. To overcome these nonlinearities, a novel correction may be used, based on an optical feedforward correction technique.
The prototype system employs off-the-shelf components, and has one Fabry Perot laser diode (FP-LD) providing two signals (via a beam splitter), for a main path and one for the error path loop. The error path signal is detected by a receiver circuit, then mixed with a reference signal to produce a 'pure' error signal, which then modulates the second FP-LD.
In contrast with previous fibre feedforward systems, where the two LD outputs are then combined in the optical fibre pre-reception, this system has to combine the signals post-reception. After the main signal and error signal are received and recombined, the non-linearities of the main path are predominantly cancelled by those present in the error path signal, leaving only the desired signal, free of non-linearities.

Chapter 1 - Introduction

This thesis presents research concerning a free space non-linear correction system, using a feedforward correction technique, with the aim of using it as part of an access link between a base station and the antennas of a digital cellular phone network.

The need for such a link is to introduce flexibility into the access network, while the correction system is needed to achieve a dynamic range of typically 80dB in such a link.

It is to be noted that rather than being a particular solution for any one format, this system aims to prove a principle that can then be adapted as required, although some key design boundaries will be formed and adhered too.

The remainder of Chapter 1 is concerned with reasons for needing remote antennas and some of the key design requirements. Chapter 2 then explains the technicalities of the non-linearities to be cancelled and the various techniques available to achieve this. The mathematical modelling of the system is contained in Chapter 3, as well as a brief explanation of the models used. Chapter 4 is concerned with the noise analysis of the system, with the practical design and set up of the system, along with the discussion the results from the practical experiments in Chapter 5. Finally, Chapter 6 has the conclusions and ideas for further work.

1.1 Background

It has always been vitally important to mankind to communicate further than the acoustic range of the human voice. However it is only comparatively recently that 'global' communications as we know them have become possible.

Previously, communications, at least in the sense of real time, was only really possible on a line of sight basis. An example of these are light towers, (used by the ancient Greeks), smoke signals (North American Indians), and flags (numerous armies and navies throughout history).

More recently the telegraph gave the first electronic long distance communication using Morse code.

It was not until the invention of the Telephone by Alexander Graham Bell in 1876 and the Radio by Guliemo Marconi in 1895, that the field of communications started to truly expand.

The telephone, from the Greek tele, 'far,' and phone, 'sound,' involves the concept of speech sounds as being a complex set of vibrations in air that is transferable to solid bodies and of the convertibility of those vibrations to electrical impulses (in conducting metals),

[1].

In the following years the telephone instrument, as modified by Thomas Watson, Emil Berliner, Thomas Edison, and others, acquired a form that has not changed fundamentally in a century. Since the invention of the transistor in 1947, bulk circuitry and other heavy hardware have been replaced by lightweight and compact microcircuitry. Advances in electronics have improved the performance of the basic design, notably analogue-to-digital conversion for transmission over digital circuits that combined with the introduction of optical fibre networks, gave rise to the long distance, high capacity and hence, inexpensive, communications world-wide. Such advances have supplemented, but not replaced, the basic telephone design. Changing little since the early years of telephone communication, the telephone instrument consists of the following functional components: a power source, a switch hook, a dialler, a ringer, a transmitter, a receiver, and an anti-sidetone circuit.

During this period, the telephone exchanges evolved from mechanical relay based systems, such as Strowger, to the modern digital switches today that can handle enormous amounts of traffic. Also, with the proliferation of data services, such as the Internet, the rise of other traffic routing protocols, such as IP (Internet Protocol), are becoming prevalent [2], [3], [4].

However, as far as the end user was concerned, there was little significant change in the telephone until the advent of the first generation analogue mobile radio transmitted phone in the early nineteen eighties. These were limited, at first, in both performance and geographical usage [4], [5]. However, as the technology improved, and hence, with price

reduced, they became more popular. After only a few years the analogue spectrum allocated for mobile radio transmission reached its capacity and was one of the reasons behind the introduction of the more sophisticated, second generation, digital mobile phone, introduced in the early nineteen nineties, giving a much higher capacity of users per unit bandwidth.

These are the phones presently available, having been available for around ten years now, and include the European GSM (Global System for Mobile communications) format that allows global roaming and limited data services.

Improvements on the GSM format are still ongoing, and will be for a few years to come. These include GPRS and EDGE [2], [3], [6], [7]. However, it is the introduction of the third generation, or UMTS, mobile (which was introduced in Europe in early 2003) that is one of the most exciting areas of communications. This allows voice, e-mail, Internet and other multimedia services from the user mobile phone terminal [2], [8], [9].

The remainder of this chapter will describe the UMTS format, being the driver behind this research, before discussing the concept of remote antennas, and the physical options available to realise them.

1.2 GSM, UMTS and 3G

The radio, in its traditional 'wireless' guise, has been steadily evolving ever since it was invented, a major step being the introduction of FM around thirty years ago and, more recently, the advent of digital radio for public broadcast [10].

However, it is in the mobile phone networks that digital radio has seen the most interest, as all second generation, or 2G, mobile phones use this technology. GSM is the European standard for 2G phones.

Apart from the great desire for extra increased bandwidth services, the other great market driver behind the development 3G and UMTS (which is a very market driven concept), has been the need for a more standardised global format, allowing seamless roaming with just one phone.

Issues of standardisation have their origins back in the 1st generation analogue phones, in which almost every different country had its own standards, especially at the air interface, making cross-country compatibility and roaming problematic.

These problems were somewhat alleviated with the introduction of the 2G phones with areas such as Europe adopting a common standard, GSM. As 2G expanded in popularity, to ever more countries, GSM became the world's leading format. However, there are still significant differences between GSM and the leading 2G formats of the Far East and North America, which stem largely from backwards compatibility with analogue networks, but also from different views of network formats and organisation.

These differences can include the form of modulation technique chosen at the radio interface, the access scheme, the multiplexing and even the type of cellular delivery network, (for example, the ill fated Iridium satellite cellular network [11]). This is not to say that satellite or any other particular solutions are not useful; they just need to be part of a larger integrated network.

One of the key drivers behind the evolution towards 3G, as previously stated, is network convergence, of both fixed and mobile technologies, with near-complete cross compatibility. This is not to say that there will not be differences in the air interface formats across the world, just that the formats will be similar enough that the handsets/receivers will be designed so able to adapt to them, thus giving the end user full global roaming, (providing they are of course willing to pay for it), while providing full backwards compatibility for 2G users in their original areas or countries.

An important aspect of the 3G development is that it is an evolution from 2G rather than a total 'hard cut' replacement. Taking the European GSM/UMTS format as an example, there are several intermediate stages involved in the progression from GSM to UMTS. A standard GSM phone can receive data at a rate of 9.6 kbps, next comes HSCSD (High Speed Circuit Switched Data) that utilises multiple channels to give a data rate of up to 64 kbps. Following this is the GPRS (General Packet Radio Service) that uses IP style packet services to offer data rates up to around 170 kbps. These formats are what drive the Wireless Access Protocol phones of today. The next GSM based standard will be EDGE (Enhanced Data rates for GSM Evolution), which will offer data rates of up to 384 kbps and is seen as the format to use for those operators who failed to acquire a UMTS

licence [2], [3]. All these services, while possibly having differing modulation schemes and some minor modifications, will use the TDMA, 900/1800 MHz GSM air interface and, will probably take the lead in data services for a number of years, due to the flexibility of the network [3]. Further details of these formats can be found in several forms of literature [8], [3], and [5].

Generic to all of these and indeed to the UMTS format will be the core to access network interface [12], that allows for network flexibility and future alterations and improvements.

The driving forces behind the overall development of UMTS and its standards are the ITU (International Telecommunications Union) and ETSI (European Telecommunications Standards Institute) [13], which consist of representatives from the communications community, along with some of the major communications manufacturers, such as Nokia and Alcatel.

Within this European research program for Advanced Communications Technologies and Services, (ACTS) [14], [15], [16], the Future Radio Multiple Access System, (FRAMES) [17], [18], project was set up to investigate and recommend the standards for the UMTS air interface.

Unlike its GSM predecessor, UMTS has to be flexible enough to offer varying data rates of 144 kbps, 384kbps, 512 kbps through to a possible 2 Mbps, depending on location, mobility (i.e. speed of travel) and of course, cost. This concept is illustrated in Fig. 1.1 below.

In order to do this, the air interface, the UTRAN (UMTS Terrestrial Radio-Access Network), is to be WCDMA (wideband code division multiple access), have a bandwidth of 5 MHz and be transmitted in the range of 1900 MHz to 2170 MHz. In addition to this, two sub-schemes of W-CDMA FDD (frequency division duplex) and TD-CDMA TDD (time division duplex) are to be used, with multiple data modulation techniques, such as QPSK and 16QAM, depending on the data rate required, location, mobility and so on. A general comparison of 2G and 3G systems can be seen in table 1.1, while an expanded summary of the differing schemes derived from the FMA, (FRAMES multiple access), which is the specification for the UTRAN, can be seen in table 1.2 below [17], and an in depth analysis of CDMA can be found in [19].

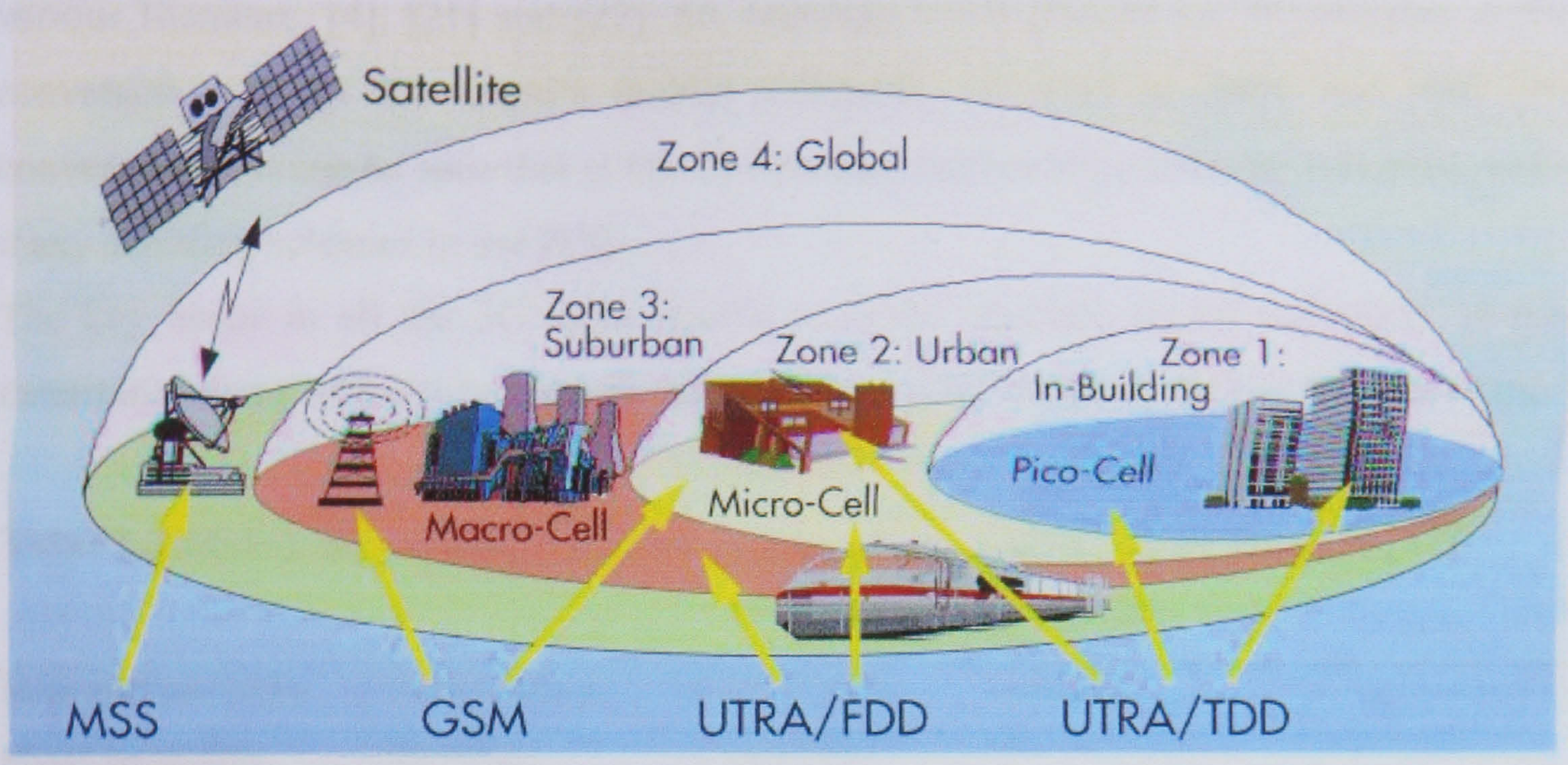


Fig. 1.1 *UMTS concept, [20]*

Table 1.1 *General comparison of 2G and 3G parameters*

	2G	3G
Bandwidth	1.25 MHz	5 MHz
Data rate	9.6 Kbps	2 Mbps
Frame length	20 ms	10 ms
Circuit & packet switched data	No	Yes
Simultaneous voice & data	No	Yes
Power control	Reverse open and fast closed, forward slow control loop	Forward and reverse open and fast closed loop (16KHz)
Handover	Soft handover	Soft handover, interfrequency handover
Multirate service	No	Yes

It is worth bearing in mind that these are just the schemes for UMTS and that other 3G standards, whilst similar, do vary the world over, summaries of which can be found in

various literature, [4], [21] and [22]. So, although a key goal of the 3G standard is the convergence of all the world's mobile standards, (as well as fixed and land line convergence), it can be seen that at the air interface, technically, there are still going to be many differing schemes in use [23].

The key factor in all the 3G radio interfaces is the multiple access technique, as this determines many of the other system characteristics [24]. Some of the key features of the

Table 1.2 *Main FMA parameters, further explanation of table can be found in [17]*

Main multiple (MA) access parameters	FMA1 without spreading	FMA1 with spreading	FMA2
MA method	TDMA	TDMA/CDMA	DS-CDMA
Duplexing method	FDD and TDD		FDD
Basic channel spacing	1.6MHz		5MHz
Carrier chip/bit rate	2.6Mb/s/5.2Mb/s	2.167Mchip/s	4.096Mchip/s
Time slot structure	16 or 64 slots/ TDMA frame	8 slots/TDMA frame	n/a
Spreading	n/a	Orthogonal, 16 chips/symbol	Spreading factor 4-256, short codes for DL/UL; long code option for UL
Multirate concept	Multislot	Multi-slot/code	Variable rate spreading & multicode
Frame length	4.615 ms		10 ms
Forward error correction codes	Convolutional codes R = ¼ to 1, puncturing/repetition		Convolutional codes; UL: R = ½; DL: R = 1/3, puncturing/repetition
Data modulation	BOQAM/QOQAM	QPSK/16QAM	QPSK
Spreading modulation	n/a	Linearised GMSK	Dual channel/Balanced QPSK or complex 4-phase spreading (UL)
Pulse shaping	Root raised cosine, roll-off = 0.35	(Linearised GMSK)	Root raised cosine, roll-off = 0.22
Detection	Coherent, based on training sequence		Coherent detection, UL: reference bit based; DL: pilot/reference bit based
Other diversity means	Frequency/Time hopping per frame/slot		Macrodiversity
Power control	Slow power control, 50 dB dynamic range (DR)		UL: open loop and fast closed loop (80dB DR); DL: fast closed loop (20 dB DR)

UTRAN format WCDMA (which is also the backbone for the 3G specification, with other techniques used to other special functions when needed) are: data rates up to 2Mbps in a 5MHz bandwidth; efficient and fast packet access; multiple parallel services; inter-frequency handover support; and built in support for future enhancements such as adaptive antennas and advanced receiver structures.

Some of the important technical benefits of this technology are: service flexibility of circuit and packet switched services; enhanced spectrum efficiency; increased coverage and capacity; good economies of scale, by adding the WCDMA access to existing networks, such as GSM; seamless global roaming; relative terminal simplicity and therefore economy compared to other technologies and multiple services per connection. An expansion of these benefits, with further explanation of how they are to be delivered, can be found in [3], [17] and [25].

Further explanation of the technical aspects of the FRAMES standards for the UTRAN, such as coding techniques, uplink/downlink specifications and specific modulation techniques can be found in many sources. However good summaries can be found in [17], [25], [26], [27] and [28].

In summary, it can be concluded that it would be far more desirable to design a system that can prove the principle and then be adapted to work with any format, rather than to design one for a particular format as mentioned above. Therefore, the only criterion of the UTRAN that has been met is the bandwidth of 5MHz and the modulation frequency of around 2GHz, although due to reasons that will be explained later, this was subsequently changed. The issue of the modulation scheme will also be considered later. This proof of principle is important because it has been stated that any radio interface should be capable of operating in any feasible frequency band that becomes available, [24].

1.3 Remote Antennas

A remote antenna is when the mobile cellular antenna is positioned remotely from its base transceiver station, BTS, rather than on the same site connected by microwave coaxial cable, which has a maximum useable length of around 10m. Usually this will

mean the BTS and antennas are connected via fibre optic cable, but can also mean via free space. The following discussion is applicable to either the fibre, or free space, techniques.

As previously stated, the mobile communications industry has expanded rapidly in recent years and this growth is due to continue, (albeit at perhaps a more consistent rate than previously thought and/or experienced, due to the global slow-down of the technology markets and the events of September 11th, 2001), with the introduction of 3G technology and the provision of new personal communications services. This is creating pressure on the existing infrastructure, and is leading to research into solutions to reduce this pressure, including remote antenna systems.

This problem becomes all the more acute in urban areas. Buildings and tunnels can impede the radio signal, causing shadowing and multipath problems. The traditional solution to this problem is to increase the amount of microcells within an area, making them smaller, often to the point of becoming picocells, so everywhere is covered. Microcells tend to have radii of between 100m to 500m, and picocells from 10m to 100m, following the configuration as shown in Fig. 1.2 below:

Fig. 1.2 *Cell distribution with frequency reuse*

Each hexagon is a frequency cell, with each number and colour representing its own allocated frequency. The antennas are usually situated around 15m above ground and their BTS. This system allows the re-use of frequencies, as long as no neighbouring cells have the same allocated frequency. With this traditional solution of smaller cells and increased frequency re-use, this also means an increase in the number of required BTS (i.e. one for each cell), which are rather complex and expensive pieces of hardware. Besides the financial cost of the BTS, there is the problem of purchasing the appropriate sites for them, which can have financial and social penalties.

A solution to this is to use small, low powered transceiver antennas in each microcell, connected to a central BTS that deals with all the signal processing. Both the antenna and the BTS use an electrical to optical conversion, i.e. a laser, to transmit to one another and visa versa, i.e. a photodiode, when receiving, with the uplink classed as antenna to BTS, and the downlink the BTS to antenna. An illustration of such a system can be seen in Fig. 1.3, below:

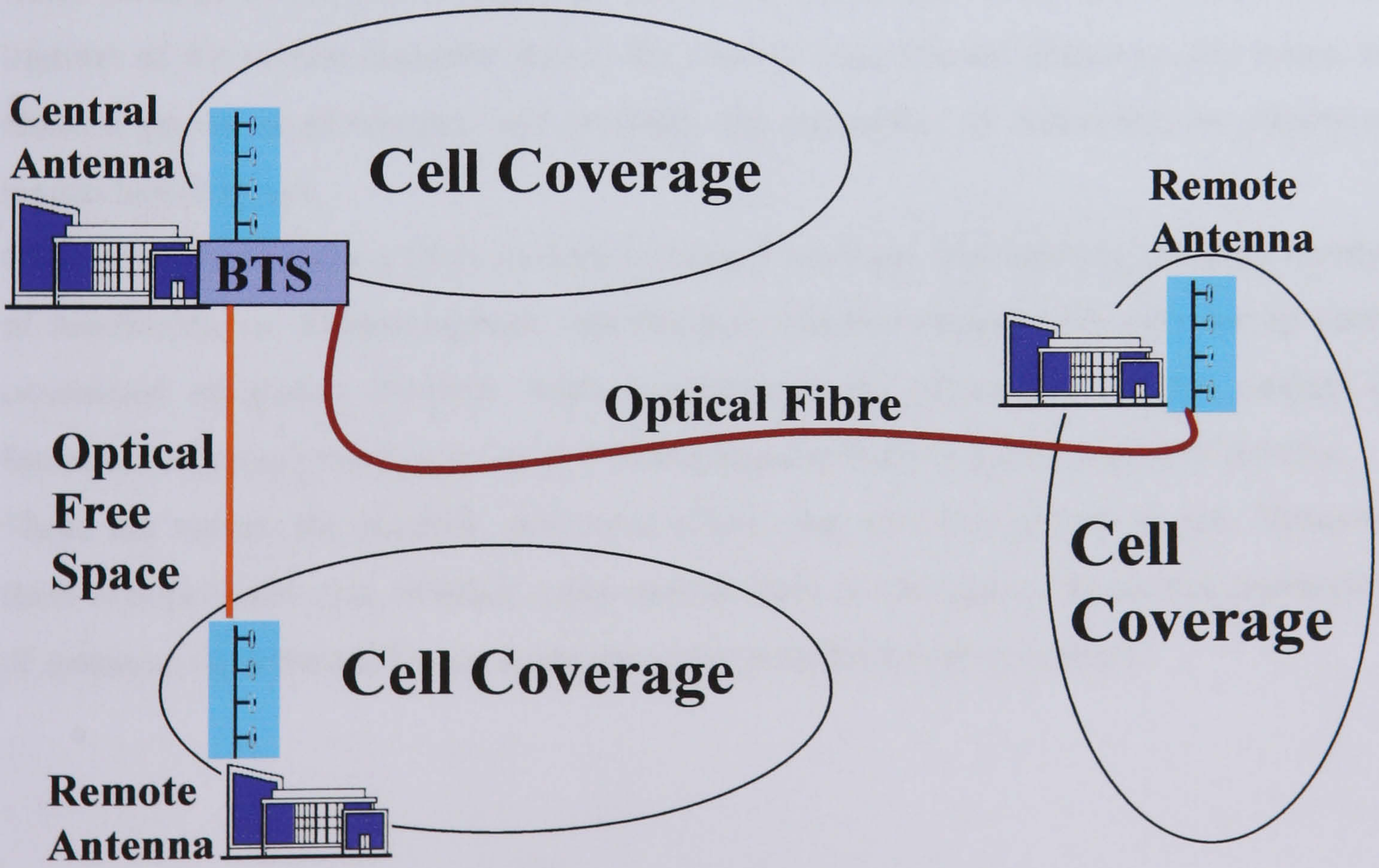


Fig. 1.3 *Principle of remote antennas, over optical fibre and free space*

Having a distributed antenna system (DAS) such as this has many benefits, most of which come from having small, low power antennas, (albeit maybe more of them).
Firstly there is the reduced physical environmental impact due to the reduction in the amount of bulky BTS needed, with the general disruption that goes with their installation. More important, though, is the fact that, as the distance between the mobile handset an antenna is reduced, so the power radiated by both can also be reduced. The first benefit of this is that more frequency re-use within microcells is possible, because interference is reduced, thus creating picocells. This increase in the number of cells and hence, number of frequency reuses, leads to, as demonstrated by the simple formula below [29], increased spectral efficiency, η_s (erlang/m^2/Hz), which is very much one of the key aims of cellular mobile network designers:

$$\eta_s = \frac{\text{number of reuses}}{\text{coverage area}} \text{ x } \frac{\text{number of channels}}{\text{bandwidth available}} \text{ x } \frac{\text{time the channel is busy}}{\text{total time of the channel}} \qquad \textbf{(1.1)}$$

Other benefits of low power antennas include the following: firstly the increased battery lifetime of the mobile handsets due to the smaller transmission distances and hence the reduced power requirements, and secondly the relaxation of perceived, or otherwise, human health issues.
Other benefits of using a DAS include increased coverage and capacity, reduced number of handovers, no electromagnetic interference, reduced design and engineering costs, centralised up-grades, dynamic radio resource/capacity allocation, and the support of future broadband applications, all of which help contribute to a high quality of service.
These are not all the benefits, and some others may not be apparent as yet. However, there is little doubt that, whether using optical fibre or free space, the remote positioning of antennas from their BTS has some attractive benefits for all concerned.

1.4 Optical Free Space and Optical Fibre

The idea of sending RF signals down an optical link is not especially new, with research into transmission of cable television, CATV, on optical fibre well-established around eight years ago [30], [31], with development still ongoing [32], [33]. However, it is only in the last three years that this technique has been seriously considered for the applications in the mobile digital phone network, whereas the use of free-space links in this role, very much the aim of this project, is an even more recent development.

For standard telephone digital communications, such as PDH and the more recent SDH trunk communications optical fibre has been in use for around three decades [34], and is the backbone to the world's telephone network. The high bandwidth capacity that fibre can carry has become further exploited now that internet traffic has become so popular. Now optical fibre is being installed into businesses and homes, to give ever higher data rates. And, in areas where fibre is not appropriate, then optical free space applications can be considered, such as the communications network set up around 'ground-zero' New York, post September the 11th [35], [36].

Firstly, it needs to be stated that these two systems actually have far more in common than they have different from each other, with many of the same properties, problems and subsequent solutions. One of the key properties is that the direct modulation of an optical light source, (in this case a laser), neglecting for a moment non-linearities and noise, makes the operation of remote antennas independent of the signal's modulation format, although care should be taken when considering the optical modulation index when certain large numbered QAM schemes are to be employed. This independence is quite a useful property, considering the variety of formats used within mobile communications, as previous sections have highlighted. Another shared characteristic is the problem of non-linearities that arise from the laser operation, and the various techniques that may be employed to correct these problems. Further explanation of the non-linearities involved, and the correction techniques available, will be found in the next chapter.

The next characteristic of concern is noise. Whilst noise is considered less of a problem compared to the non-linearities in this application, it cannot be ignored, and, therefore

there is a difference in the two systems. While the two systems will have similar components in the transmission units and receiving units, the difference is, of course, in the transmission medium. In fact, both media, fibre and air, can be fairly noise free, providing suitable technology is employed. The difference comes in the form of signal attenuation, the susceptibility to the noise present and the ambient illumination in the free space environment.

Optical fibre has very low attenuation, and, in the case of SDH digital (on-off keyed) trunk lines, it is used to transmit signals of around 2Gbps over distances of at least 100km without regeneration [37], [38]. Moreover, soliton-pulse WDM transmission is reaching up to 1300km without regeneration, by using sections of dispersion-compensation fibre, at very high data rates [39].

Although SDH equipment tends to be of the highest specification, standard optical fibre has an attenuation of less than $1dBkm^{-1}$ [34]. Free space communication has more variables, and though transmission attenuation can be low for a short link, certain atmospheric and climatic conditions, such as fog, can have a significant impact on the transmission distance. This topic is covered in greater detail in Chapter 2. Just transmitting a more powerful signal is not a solution, due to eye safety issues, with the maximum power allowed at around 1mW for a typical beam, (or $5.76Wm^{-2}$), depending on the wavelength of the light. This subject is covered in greater depth in Chapter 5. Optical fibre systems do suffer a coupling loss at each end of the fibre; however, this is a minor loss and is often ignored in terms of overall system performance. An in depth analysis of optical fibre performance when used in a digital cellular application can be found in the ROBIN Project final report, [40].

Another significant disadvantage of a free space system is that it has to have 'line of sight' between the transmitter and receiver, which is clearly not the case with a fibre system.

An advantage a free space link does have is that it does not suffer from fibre dispersion. Chromatic dispersion is when light of differing wavelengths propagate along the fibre at different speeds due to the properties of the optical fibre. This is generally overcome by using single mode fibres and lasers with a small line width. However, some of the non-linear correction systems that can be used have more than the one signal at differing

wavelengths, leading to various dispersion induced problems, which will be clarified in the next chapter. However, it can be seen that, as free space effectively does not suffer from this form of dispersion, this potential problem is avoided.

Another significant advantage is that because there is no man made transmission medium, there are none of the problems of installing such a medium, such as the digging up of roads, and the expense associated with such requirements. As part of the driving force behind remote antennas is to lessen the social and environmental impact of cellular networks, the extra flexibility that free space can offer, avoiding such disruption, is appealing.

There are many more comparisons that might be made between the two technologies. However, these are largely irrelevant, as it is unlikely that the two will be 'competing' against each another. For the most part, the remote positioning of antennas will be undertaken using optical fibre, especially when distances are in excess of a few hundred metres, after which the attenuation of optical free space transmission become problematic. However, there will be circumstances where a fibre system is not practical, too expensive, too slow to install, or just a less attractive option than using free space, due to the latter's extra flexibility. Such situations could be indoor communications, such as within shopping malls, and outdoor communications, such as in dense urban areas, where shadowing causes coverage problems. Here, a link between two buildings could provide a quick and inexpensive solution, even if it was only temporary while a more permanent solution was being realised. Also, rural areas and natural features, such as rivers, could be negotiated in much the same way, as a permanent or temporary solution.

In summary, it can be seen that the need for remote antennas in the modern digital cellular networks is very real and that several solutions are available, each of which will be valid, depending on the needs of a particular situation.

Chapter 2 – The Non-Linearities present and the techniques available to correct them.

One of the major concerns with any communications system is the issue of noise and interference. The general definition of what makes up noise, can encompass any signal present at an output other than the one desired. Sources of noise are numerous, ranging from the conventional sources, such as thermal (or Johnson) and shot noise. Though unwanted, interference from other signals, whether it is in the form of harmonic distortion or intermodulation distortion, arising from non-linearities of various components, are not classed as noise.

Therefore, noise is thought of as those sources such as thermal and shot noise, whereas those associated with non-linearities are usually classed as interference or distortion and require a signal input. The subjects of noise and non-linearities in wireless and fibre communications systems are widely covered in numerous texts [34], [41], [42], [43], [44] and [45], therefore, the following six sections are intended to serve as a summary. The penultimate section will be an introduction into the various linearisation techniques available within communications, with the final section introducing a novel free space implementation of one of the linearisation techniques discussed.

2.1 Link design specifications

The noise establishes a noise floor that determines the minimum level for an RF signal that is detectable for a given link. The difference between this and the desired carrier power for a given channel, for its associated channel bandwidth, is the carrier-to-noise ratio, or CNR. With careful circuit/system design, this noise floor can be kept to a minimum. However, maintaining the required CNR over a certain distance, whether the system is free space or fibre, will often require an increased input signal. This, in laser

diode systems, can lead to increased non-linearities, which in turn tend to be the limiting factor for system performance once they appear above the noise floor. Therefore, if the non-linearities can be cancelled, the system can be driven harder, leading to greater distances of transmission and hence, greater utility of use.

It is worth noting that other components in the communications link will exhibit non-linear behaviour, such as the photodiodes, PDs, and any amplifiers used. However, it is usually the case that the non-linearities produced in the PDs and amplifiers (providing careful selection) will be minor to those produced by the laser diode itself [46], [47], [48].

The parameter of performance that is of most importance is the Dynamic Range, DR, and is defined as the difference between the noise floor and the maximum possible input signal level. The DR has two sub-parameters, the 1dB Dynamic Range or DR_{1dB}, and the Spurious Free Dynamic Range or SFDR.

The DR_{1dB} is the simplest limitation and is defined as follows. When the output signal is reduced by 1dB from that predicted by the gain of the link due to clipping of the input, the 1dB compression point, P_{1dB} has been reached. The DR_{1dB} is then the difference between the input level that gives the P_{1dB} and the noise floor. In dBs, the relation is as follows (and is further illustrated in figure 2.1 below) [41];

$$DR_{1dB} = P_{1dB} - Noise_{channel,dBm} \tag{2.1}$$

The drawback of the DR_{1dB} measurement is that it only defines the link performance for a single frequency input.

For multiple frequency inputs, the SFDR gives a far more precise evaluation of link performance. SFDR is defined, in the two tone case, as the difference between the level of the two (equal) tones and the noise floor, where the intermodulation products are equal to, or below the noise floor. In dBs the SFDR can be expressed as [41];

$$SFDR = 2/3(IIP3_{dBm} - Noise_{channel,dBm}) \tag{2.2}$$

The term IIP3 is the input third order intercept point (third order intercept, or input TOI) and is often used to characterise the 3rd order distortion of an amplifier. When two equal

sinusoidal tones modulate a laser, third order distortion products are generated at $2f_1 - f_2$ and $2f_2 - f_1$ and are the intermodulation distortion products, or IMD_3, (the subject of distortion will be expanded upon later). When these products are expressed in relative terms, in dB's and plotted against the input or output power levels, it can be seen that they have a gradient of three, compared to a gradient of one for the fundamental. The fundamental and IMD_3 slopes are then extrapolated until they intercept. The input power corresponding to this intercept is the IIP3, as can be seen in Fig. 2.1 below. It is important to note that the IIP3 cannot actually be measured, as driving a device so the IMD_3 product is equal to the fundamental would exceed the device's maximum power rating. Instead, the IIP3 can be calculated using the following formula [41].

$$S/I_{2tone,dBc} = 2(IP3_{dBm} - S_{in,dBm}) \qquad \textbf{(2.3)}$$

Where $S_{in,dBm}$ is the input signal and $S/I_{2tone,dBc}$ is the difference between the input signal and the level of the IMD_3 product at one of the frequencies of $2f_1 - f_2$ or $2f_2 - f_1$. In practise, the IIP3 is often quoted for a device. For example, if a laser unit had an IIP3 of 35dBm and an input of 0dBm, then the S/I_{2tone} would be 70dBc and the IMD_3 would be at –70dBm. In lasers, both the 1dB compression point and the IIP3 point are dependent on frequency and bias current [41].

In an RF radio link for mobile phone networks, it is generally assumed that the up-link, that is from the phone/mobile terminal to the BTS, will need more stringent requirements than the down link. Although many digital cellular systems only require an SFDR, (or approximate CNR), of as low as 30dB, the fact that the carrier amplitudes of the uplink signals can vary in strength by around 40dB+, and the weakest signal needs to be accommodated. This means an SFDR of 70dB+ is required. This is an approximate benchmark for the uplink of a GSM system, [42]. It is thought that UMTS may not be as strict this, although, (as stated in the previous chapter), from the viewpoint of this work, specific dynamic range targets are less important than the principle of improvement in dynamic range, and performance in general, by the use of linearisation techniques. Therefore, although specific numbers for SFDR will be given due attention, it is not the

aim of this project to satisfy any one specific system requirement, but offer a possible solution to many.

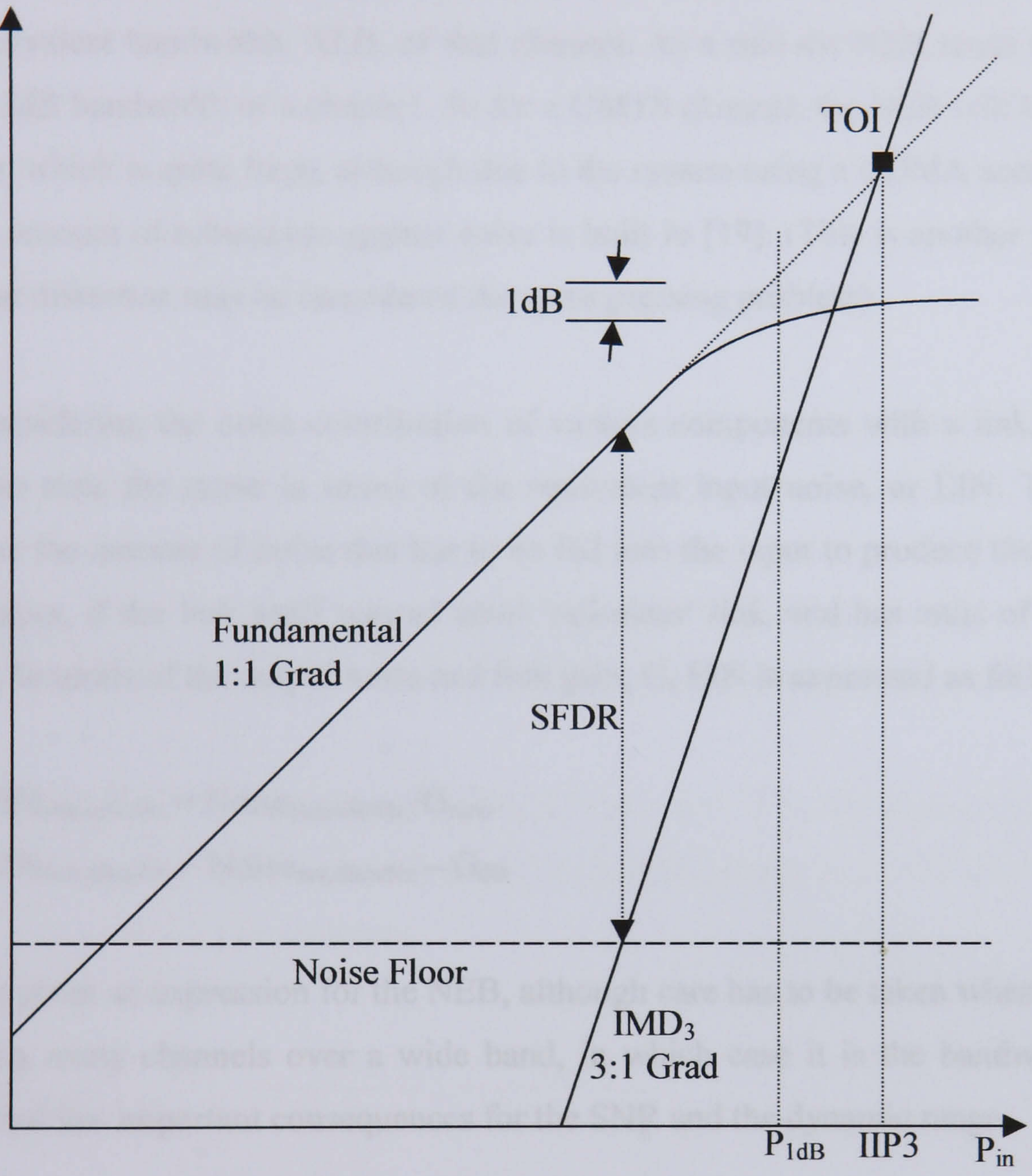


Fig. 2.1 *Spurious Free Dynamic Range*

2.2 Noise

Although it has been stated previously that in this work noise is of less concern than non-linear distortion, there are still some important parameters worth noting. A noise source is one that is effectively independent of the modulation frequency of an input RF signal and

which affects the sensitivity of the system. Noise in laser systems has four main sources; thermal noise, flicker noise, shot noise and relative intensity noise, RIN. The sources of noise are generally classed as white, and distributed evenly across the frequency spectrum. Therefore, the noise present in any one channel will be dependent upon the noise equivalent bandwidth, NEB, of that channel. As a rule the NEB tends to be larger than the 3dB bandwidth of a channel. So for a UMTS channel, the NEB will be in excess of 5MHz, which is quite large, although due to the system using a CDMA access scheme, a certain amount of robustness against noise is built in [19]. (This is another reason why non-linear distortion may be considered the more pressing problem).

When considering the noise contribution of various components with a link, it is often simpler to state the noise in terms of the equivalent input noise, or EIN. This can be defined as the amount of noise that has to be fed into the input to produce the noise seen at the output, if the link itself was an ideal 'noiseless' link, and has units of mW/Hz or dBm/Hz. In terms of the output noise and link gain, G, EIN is expressed as follows;

$$EIN_{link,mW/Hz} = Noise_{out,mW/Hz} / G_{ratio}$$

$$\text{or} \qquad EIN_{link,dBm/Hz} = Noise_{out,dBm/Hz} - G_{dB} \qquad \textbf{(2.4)}$$

This then gives an expression for the NEB, although care has to be taken when a link may be passing many channels over a wide band, in which case it is the bandwidth of the channel that has important consequences for the SNR and the dynamic range.

$$Noise_{channel} = EIN_{dBm/Hz} + BW_{dBHz} \qquad \textbf{(2.5)}$$

This expression can now be used in equations (2.1) and (2.2) if so desired.

An alternative measure of noise, is that of the noise figure, or NF, which for a two port network is defined in as;

$$NF = \frac{\text{SNR input}}{\text{SNR output}} = \frac{S_{in}/N_{in}}{S_{out}/N_{out}} \qquad \textbf{(2.6)}$$

Now, if the two port network has gain, G, then $S_{out} = GS_{in}$, (although this does not follow for the noise, as there is noise added by the network). Substituting for S_{out} into (2.6) gives;

$$NF = \frac{S_{in}/N_{in}}{GS_{in}/N_{out}} = \frac{N_{out}}{GN_{in}} \qquad \textbf{(2.7)}$$

Given the further definition that the input noise can based on an ideal device, passively terminated at the standard temperature of 290K, T_o, N_{in} is now given as;

$$N_{in} = kT_oB \qquad \text{which leads to } NF = \frac{N_{out}}{GkT_oB} \qquad \textbf{(2.8)}$$

where k is the Boltzmann constant (1.38×10^{-23}J/K) and B is the bandwidth. Once the noise figure for a component has been acquired, Friis' formula can be used to cascade individual elements within a system;

$$NF_{Total} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1G_2} + \ldots + \frac{F_n - 1}{G_1G_2 \ldots G_{n-1}} \qquad \textbf{(2.9)}$$

where n is the number of individual elements in the system. (It should be noted that equation (2.9) assumes the each amplifier has the same noise bandwidth).

The noise figure can also be related to the EIN, over a unit bandwidth by [41];

$$NF = 10 \log (EIN_{mW/Hz}/kT_o)$$

$$\text{or} \qquad NF = EIN_{dBm/Hz} - 174\text{dBm/Hz} \qquad \textbf{(2.10)}$$

Other definitions, such as noise temperature can be found in numerous texts [41], [43], [45] but will not be expanded upon here.

As previously stated, there are four main sources of noise and these can be related to the various components in the communications link. Thermal and flicker noise are related to the amplifiers the link, whether they are pre transmission or post reception. Shot noise is generated in the photodiodes, and RIN is associated with the laser.

Taking the case of the noise generated within the amplifier first. It can be seen that flicker noise, (also known as 1/f noise and not regarded as white), in which a large number of physical phenomena combine to exhibit a noise power that varies inversely with frequency, while significant at low frequencies, has little impact beyond 1MHz. Therefore, for this study, it has been ignored.

In depth analysis of thermal noise is widely available, so the following is just a brief summary. Thermal noise is caused by the thermal agitation of bound charges, resulting in random voltage fluctuations. The rms value for the thermal noise voltage, when generated in a resistance for mathematical convenience, is;

$$V_{th} = (4kTBR)^{1/2} \qquad \textbf{(2.11)}$$

where k is the Boltzmann constant, T is the absolute temperature, B is the bandwidth and R is the (nominally matched) resistance. The noise power, from $P = V^2/R$, can then shown to equal kT over a unit bandwidth, which at a room temperature of 290K, gives –174dBm/Hz . Referring this back to the input then gives;

$$EIN_{th,mW/Hz} = Noise_{th.out,mW/Hz}/G_{link} \qquad \textbf{(2.12)}$$

Shot noise arises form the fluctuation in the number of electrons emitted from a source and, in terms of the photodiode, is the result of the randomness in the arrival of the received photons. With the photodiode viewed as a current generator, the shot noise generated is;

$$I_{sh} = (2qIB)^{1/2} \quad \textbf{(2.13)}$$

Where q is the electron charge, $1.6x10^{-19}C$, I is the direct current and B the bandwidth. It can subsequently shown [42] that for a unit bandwidth, the EIN of shot noise is;

$$EIN_{sh,mW/Hz} = \frac{2qP_{laser}R_{in}L_{opt}}{\eta_{Tx}^2 \Gamma_{pd}} \quad \textbf{(2.14)}$$

where P_{laser} is the power from the laser diode (LD), R_{in} is the matched input resistance of the LD, L_{opt} is the optical loss of the link, η_{Tx} is the efficiency of the LD and Γ_{pd} is the responsivity of the PD. From equation (2.14) it can be seen the shot noise is dependent upon the optical loss, which as stated in the previous chapter, may be far higher for free-space than for fibre links (depending on atmospheric and weather conditions). Therefore, it can be seen that there will be a maximum distance for the link until the noise level becomes intolerable, even in the absence of other noise sources.

Laser noise, usually entitled relative intensity noise, or RIN, arises from random fluctuations in the level of the output around an average value. These fluctuations are caused by temperature changes, acoustic disturbances and the influence of quantum mechanics [43], [45]. These fluctuations have two forms, the intensity of the light, and the frequency of the transmission signal. The frequency fluctuation only really causes a problem in fibre systems, where it degrades the signal as it passes down the length fibre due to chromatic dispersion (the fibre characteristic of different frequencies travelling at different speeds).

Measured directly at the transmitter, RIN is defined as the ratio of the mean square amplitude of the noise fluctuations to the square of the optical power, as seen below;

$$RIN = \Delta P^2 / P_o^2 \quad \textbf{(2.15)}$$

This value is linked to the EIN by equation (2.16) [41] in which;

$$EIN_{LD,mW/Hz} = RIN(I_{dc} - I_{th})^2 R_{in} \left(\frac{M_{dc}}{\eta_{Tx}} \right)^2 \quad \textbf{(2.16)}$$

where I_{dc} and I_{th} are the bias and threshold currents of the LD respectively and M_{dc} is the DC modulation gain of the LD.

Thus, the three EINs can be added together to give the link EIN referred to in equation (2.4), as shown below;

$$EIN_{link,mW/Hz} = EIN_{LD,mW/Hz} + EIN_{sh,mW/Hz} + EIN_{th,mW/Hz} \quad \textbf{(2.17)}$$

While the both the laser and thermal noise are in themselves relatively constant, it can be seen that as some thermal noise is referred back over the link from the receiver amplifier via the G_{link} term, which in turn is proportional to $1/L_{opt}^2$, that as the optical losses increase, then the thermal noise becomes the dominant noise source, [41]. Hence, the choice of low noise post reception amplifiers are important to the system performance and will enable the system to be designed to tackle the more pressing issue of non-linear distortion.

2.3 Properties of the Air Interface

As with all communications links, the properties of the transmission medium have to be taken into account if an accurate picture of the system's performance is to be made.

For an optical free space link this medium is of course air, and the following is a summary of how the air can affect link performance in the presence of various atmospheric conditions.

The first point to make is that, much like in optical fibre communications, the air is not seen as a noise source, rather, it is regarded as an attenuator, introducing loss to the link. Hence its impact upon a link is related to the minimum noise margin of the link, i.e. the CNR, with the distance between transmitter and receiver the ultimate limiting factor.

As the LDs used in this work have a wavelength in the IR region, the following analysis will concentrate on near IR, or approximately 600nm to 1600nm. This range will cover the vast majority of multi quantum well (MQW) LDs that are currently used in optical communications.

There are three atmospheric effects that can degrade the signal quality, namely:

1. Absorption
2. Scintillation
3. Scattering.

Absorption is primarily caused by the water vapour, H_2O, carbon dioxide, CO_2 and ozone content of the air. The amount of absorption depends upon the wavelength of the light to be transmitted. Therefore, care must be taken to pick a wavelength with low absorption. Fortunately for humans and animals alike, one of the high transmission bands is the visible spectrum. However, near IR and IR are also largely unaffected by absorption, as can be seen below in Figure 2.2 below [49].

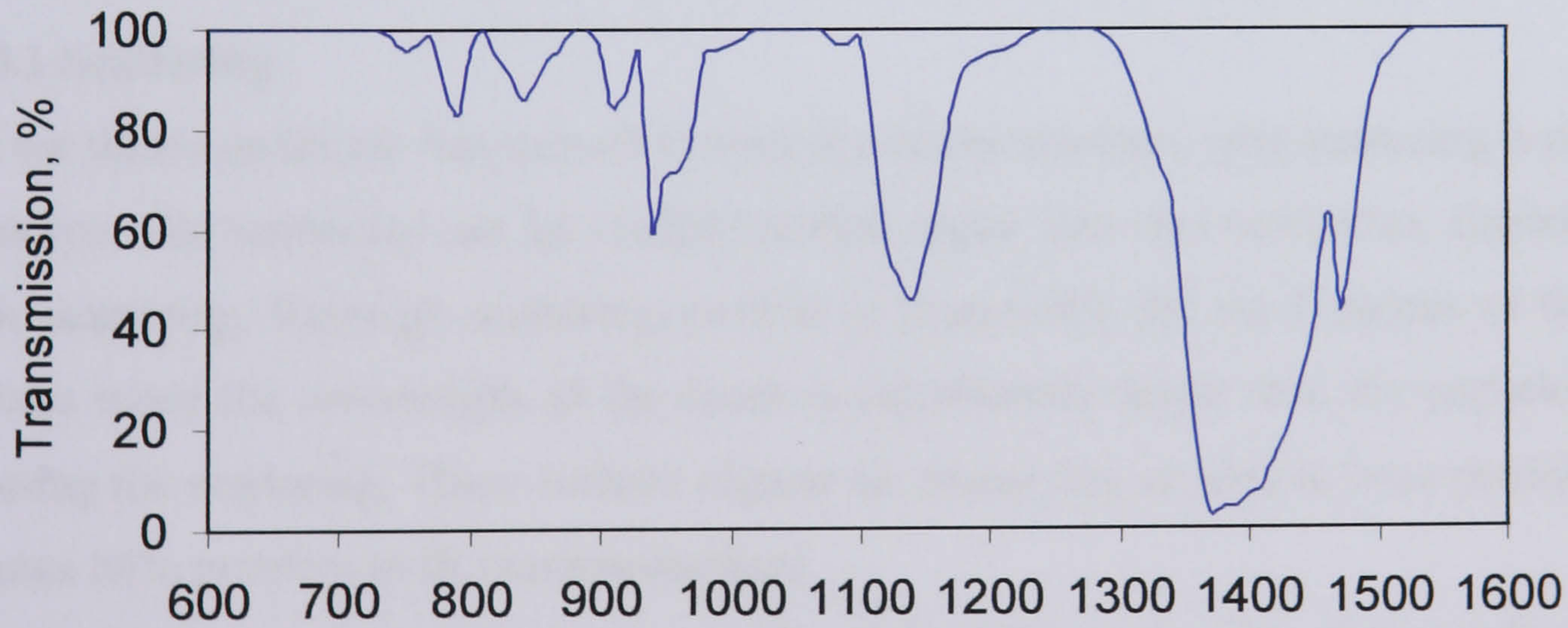


Figure 2.2 *Atmospheric transmission of IR Spectrum*

Scintillation of a transmission beam is a consequence of an effect called shimmer or atmospheric turbulence. It is the effect of when the refractive index of the air is dependent upon its temperature. When this refraction index is continually altered along the transmission path, the beam can be deviated leading to low frequency variations in amplitude and phase of the signal, also known as fading. If this fading is extreme enough to go below threshold values, errors will occur in the signal.

How far the beam is deviated is a function of the relative sizes of the beam diameter, d, and the temperature inhomogeneity dimension, l, [49]. If $d/l << l$, the effect is to deviate the beam as a whole, where if $d/l >> l$, small portions of the beam are independently defracted and distortion of the beam phase fronts can occur.

Shimmer can be best described as the effect seen when looking along a long hot piece of road, (the darker the asphalt the better). Overall, turbulence is a function of the time of day, cloud cover and wind. Simple techniques such as frequency modulation can reduce the effects of shimmer, as can automatic gain control, and adaptive optics in the receiver unit. However, the best technique to minimise beam scintillation is to position the equipment carefully as so to avoid the problem, i.e. having the system several meters above the ground and any obstacles, especially when crossing such things as black tar roofs.

2.3.1 Scattering

Of the three conditions that can affect optical communications, only scattering is of major concern. The scattering can be roughly broken down into two categories, Rayleigh and Mie scattering. Rayleigh scattering, (which is responsible for the blueness of the sky), occurs when the wavelength of the beam is significantly larger than the particle that is causing the scattering. These include regular air molecules, as well as haze particles, and causes little problem to IR communications.

The problems occur when the beam wavelength is in the same order of size as the particle radius. Such a phenomenon as fog has this property, and it is fog that is seen as the major factor/difficulty to be overcome in free space link design. The Mie scattering coefficient is a complex function of particle size and shape, refractive index, scattering angle and wavelength [49], and is usually described as:

$$\alpha = (3.91/V)(\lambda/550nm)^{-q} \quad \textbf{(2.18)}$$

where α is the Mie scattering coefficient, V is the visibility in km, λ the beam wavelength and q the size distribution of the scattering particles. Though more sophisticated expressions for q have been developed [50], (and are commented upon later), in general $q = 0.585V^{1/3}$.

An official value for the visibility, V, from extremely clear to haze to thin fog to dense (also classed as UK) fog, can be acquired from many weather monitoring services, with those taken at airports being the most closely observed.

Once a value for α is calculated, it can be inserted into Beer-Lambert's law below to find the transmittance:

$$T(z) = P(z)/P(0) = e^{-\alpha z} \quad \textbf{(2.19)}$$

Where T(z) is the transmittance at range z, P(z) is the laser power at z and P(0) is the laser power at the transmission source.

As stated above, the visibility, V, is broken down into internationally recognised codes, depending upon the visibility range, and hence, the attenuation can be calculated. This is shown in table 2.1 below.

As can been seen from table 2.1, for very clear conditions, free space communications will suffer from approximately the same loss as for optical fibre. Even for moderate fog conditions where visibility is reduced for 500m, the attenuation is still only 2.82dB per 100m. So, for a short haul free space link of 250m, the loss would be around 7dB. Providing the receiving equipment has a suitably low noise characteristic, only the worst fog conditions will cause any problems. Any non-linearities present in the transmitted signal (and indeed any noise) will suffer the same loss as the desired signal and hence, can be treated as part of the whole transmission signal, as far as this section is concerned. From Equations 18 and 19, it can also be seen that there appears to be a benefit in using the higher wavelength devices in the presence of fog, as they would suffer from slightly less attenuation.

Table 2.1 *Visibility and attenuation for different degrees of haze and fog*

Visibility Code	Visual Range	Loss, in dB/km
Dense Fog	40 – 70 m	392 – 220
Thick Fog	70 – 250 m	220 – 58
Moderate Fog	250 – 500 m	58 – 28.2
Light Fog	500 – 1000 m	28.2 – 13.4
Thin Fog	1 – 2 km	13.4 – 6.3
Haze	2 – 4 km	6.3 – 2.9
Light Haze	4 – 10 km	2.9 – 1.03
Clear	10 – 20 km	1.03 – 0.45
Very Clear	20 – 50 km	0.45 – 0.14
Extremely Clear	50 – 150 km	0.14 – 0.03

When water droplets increase in size to millimetre dimensions and above, it tends to fall as rain, which results in both scattering and absorption. The attenuation increases with the rate of rainfall but also depends on the size distribution of the raindrop, making generalisation problematic.

However, rainfall, even of a heavy nature, does not cause anything like the attenuation of fog. This is due to the radius of the raindrop being approximately a thousand times larger than that of a fog droplet, making the attenuation from rain approximately a hundred times less than for fog. Although more water falls in rain, the droplets further spaced apart and the more coarse droplets do not affect transmission as much.

Studies by Chu and Hogg of Bell Labs in 1968 [49] showed that for extremely heavy rainfall of 100mm/hr, i.e. extreme monsoon conditions, that the attenuation was still only approximately 18dB/km, the same as for light fog.

Both the attenuation and forward scattering properties of snow appear to be between those of fog and rain [49] and number of investigations into relationships between snowfall and attenuation have been undertaken. However it is considered that if a system

is designed to cope with light fog (1km to 0.5km) or worse, snow will not cause too much problem.
Another point worth noting is, where fog is not allowed to stand still, due to street traffic or thermals from buildings, the attenuation is reduced slightly.

Once the receiver noise floor for a system is calculated or measured, then using the above calculations a maximum distance for an optical communications link can also be calculated, given the minimum required link margin. There are a few innovative solutions to the problem of dense fog, such as multipath transmission at multiple wavelengths, and the advanced 'microlens' concept from Blue Sky Research™ [51]. However, these are far from inexpensive options and the technology behind them is closely guarded. The most practical and readily employed solution is the clever and sensible design of the optical communications network, especially in cities or areas where fog is a common occurrence.

As further, recent research [50], has shown that the above calculations can be made even more accurate, by using a more advanced value for the size distribution of the scattering particles that has been proposed, where the value q, of equations (2.18) and (2.19), is now:

= 1.6	for	V>50km
= 1.3	for	6km<V<50km
= 0.16V+0.34	for	1km<V<6km
= V-0.5	for	0.5km<V<1km
=0	for	V<0.5km

This indicates that in heavy fog situations, the attenuation becomes independent of LD wavelength and the scattering can now be defined as non-selective or geometrical scattering.
To conclude, in locations where fog may cause problems, care needs to be taken in the positioning of free space systems, and within the systems, it is especially important to have very low noise receiving devices.

2.4 Non-linear Laser Diode Distortion

As stated previously in this chapter, as the non-linearities produced in the other system components are small compared to those produced in the LD. Therefore, it is only LD non-linear distortion that is addressed here.

LDs produce three types of distortion: (a) static, (b) dynamic and (c) clipping distortion.

The LD that are used are Multiquantum Well, or MQW lasers, have a higher differential gain than conventional bulk lasers, which gives them a higher modulation bandwidth and a greater speed of operation. In these MQW LD there are many sources generating non-linearities, with some of the more prevalent, listed below:

1. Longitudinal spatial hole burning, or LSHB.
2. Gain compression.
3. Leakage current.
4. Relaxation oscillation.
5. Carrier transport in a QW active region.

Static LD distortion is the result of the, above threshold, non-linear light verses current (L-I) curve characteristic, shown in Fig. 2.3 below.

The effects that contribute to this non-ideal response are; spontaneous emission, although this can be ignored once above the threshold region, leakage current, gain compression and LSHB. The leakage current is that which flows outside of the active layer of a laser [52] and becomes significant in high bias current conditions, regardless of the modulation frequency. Gain compression corresponds to the reduction in gain at the higher output powers, and is the consequence of phenomena which inhibit carrier replenishment within the LD and corresponds to the wavelength of the output light. Several phenomena can cause gain compression, including spatial hole burning. LSHB is a result of the fluctuation in the carrier density along the laser cavity that results in a variation in the mode loss, or gain, with optical power.

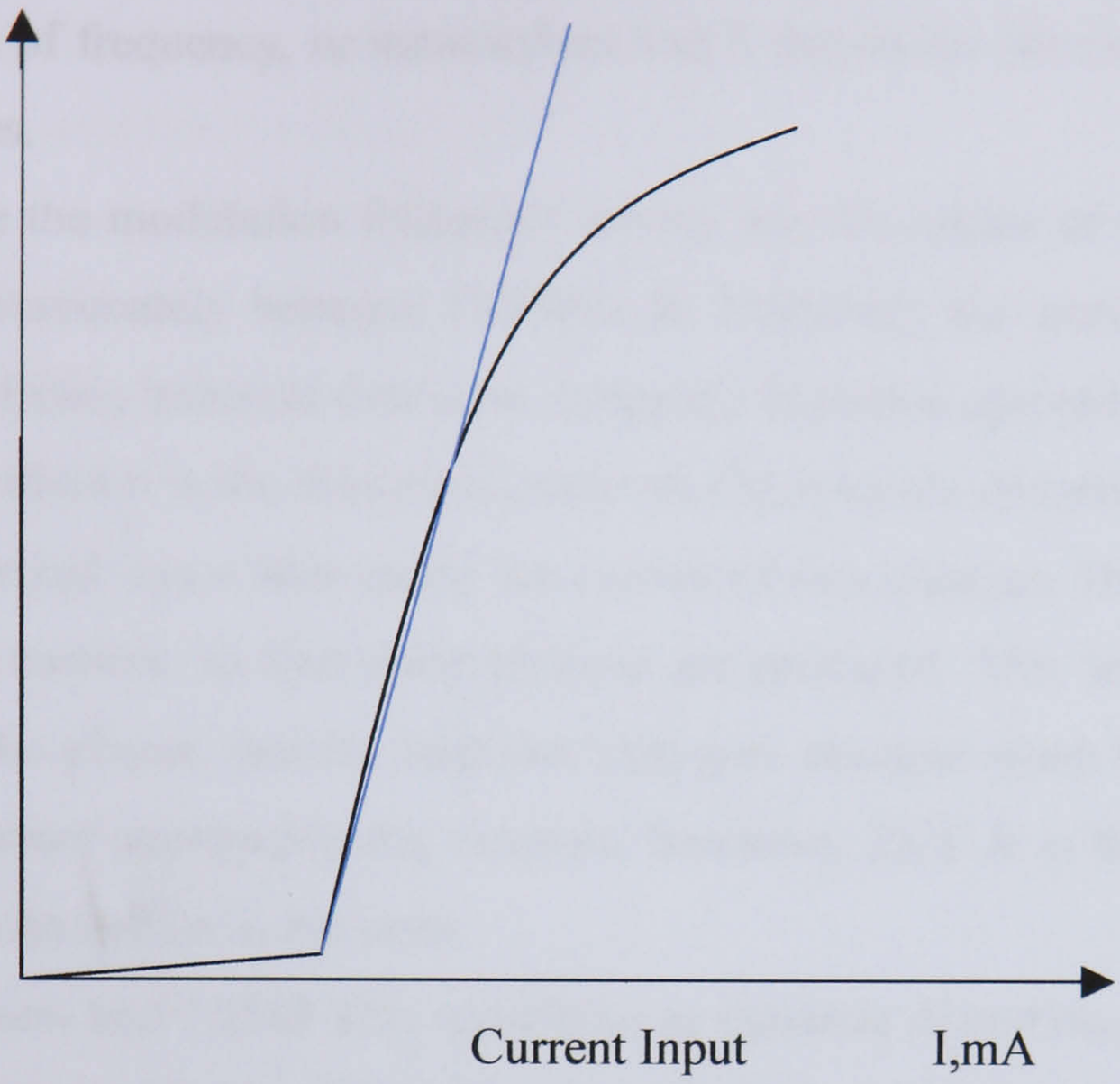


Fig. 2.3 *Light verses current characteristic, with the ideal in light blue*

As stated above,the distortion resulting from the non-ideal L-I curve is known as; (a) static distortion, and can be accurately described by the expansion of the Taylor's (power) series, as seen below;

$$v_o = a_1 v_i + a_2 v_i^2 + a_3 v_i^3 + a_4 v_i^4 + a_5 v_i^5 + \ldots. \qquad \textbf{(2.20)}$$

where v_I and v_o are small, time varying quantities representing the RF input and output signals, and the a_n are scalar coefficients. For a two tone input case;

$$v_i(t) = v \sin(\omega_1 t) + v \sin(\omega_2 t) \qquad \textbf{(2.21)}$$

Static distortion, or more accurately, weak static distortion, as has been described above, is independent of frequency, or memoryless and it dominates the distortion of the LD at low frequencies.

However, once the modulation frequency moves into the region of a few hundred MHz, (which is, approximately between 100MHz to 200MHz), the distortion caused by the relaxation oscillation becomes dominant (clipping distortion ignored).

Relaxation oscillation is the dominant cause of; (b) dynamic distortion, and occurs when the carriers injected into a laser cavity are converted into photons. This stimulates further recombination carriers, so that more photons are produced. This interaction produces a resonance in the photon density response that gets stronger when the frequency of the modulation current approaches the resonant frequency [52]. It is this phenomenon that causes the non-linearities in the laser.

Gain compression and LSHB also contribute to dynamic distortion, which dominates at higher frequencies and being clearly frequency dependent, it has memory.

Because of this memory, the simple Taylor's series is no longer entirely accurate and more sophisticated models based on the rate equations of the laser need to be employed. The Volterra series [53] is one such technique and generally thought the most accurate, and along with another technique known as Perturbation, will be used to analyse the correction system to be used, in the next chapter. The reason these techniques will be employed in this work is that having a modulation frequency of around 2GHz, means that static distortion can be effectively ignored, and the dynamic distortion is likely to dominate. However, due to its simplicity, the Taylor's series is still often used to model laser behaviour even at high frequencies, and therefore, will also be used to model the system to give a comparison to the more sophisticated techniques.

The form that the dynamic distortion will take is that of harmonic and intermodulation products. In the frequency domain, harmonic signals are those that appear at multiples of the input signal frequency. For example, if the laser is modulated at 2GHz, then harmonics will appear at 4, 6, 8, and 10.... GHz and so on, albeit decreasing in amplitude with the increasing order. In small signal analysis it is generally thought that only the 2^{nd} and 3^{rd} order distortion products will be of sufficient magnitude to cause a problem,

therefore only the harmonics at 4GHz (2^{nd} order or 2HD) and 6GHz (3^{rd} order or 3HD) would be of any interest.

Intermodulation products are those that result from the combination of two or more input signals, i.e. for a two tone signal, f_a and f_b, the 2^{nd} order intermodulation products, ID2, would be at $f_a \pm f_b$ and $f_b \pm f_a$. Similarly, the 3^{rd} order intermodulation products, IMD_3 for a two tone case, would be at $2f_a \pm f_b$ and $2f_b \pm f_a$. Intermodulation products will also appear at the fundamental frequencies of f_a and f_b.

The sum of the 2^{nd} order distortion products within a channel is also known as the 'composite 2^{nd} order' distortion, or CSO. The CSO is the ratio of the carrier power to that present for one 2HD plus multiple ID2 for that channel. Similarly, the sum of the 3^{rd} order intermodulation products for a channel is known as the 'composite triple beat', or CTB and is defined as the ratio between the carrier power and the power from one 3HD plus all IMD products for a specific channel.

In broadband systems, such as those delivering cable television (CATV), great care has to be given to harmonics produced, and especially CSO products, due to the chances of them falling into other transmission channels. Digital cellular systems are however relatively narrow band, with UMTS frequencies ranging from approximately 1900MHz to 2100MHz. Therefore, only the odd ordered intermodulation products are of great concern, as all other distortion products will fall outside of the frequency range of interest. This is shown in Fig. 2.4 below, with the distortions produced from 3^{rd} and 5^{th} order products shown for a two tone test, although in practise, as previously stated for small signal systems, only the IMD_3 products tends to be large enough to cause real concern. It is the problem of these 3^{rd} order intermodulation products that this work is addressing.

The third form of distortion originally listed is that of; (c) clipping which can represent a fundamental limit of a LD linearity and occurs when the driving modulating current of the LD encroaches on the threshold, briefly turning the laser off. Clipping is an example of a strong static nonlinearity and can be especially troublesome in AM/M-QAM hybrid lightwave video transmission systems, where it can seriously impair the digital signal BER [54]. The topic of clipping is a significant field in its own right and hence will

not be given too much consideration here, though it is worth noting that it has little effect on modulation schemes with 'analogue' characteristics, such as QPSK. However, as stated in the previous chapter, when large numbered QAM schemes are involved, i.e. QAM 64, care in the choice of bias current and optical modulation index (OMI) is needed. Though a large OMI is desirable for good CNR characteristics, there should not be any need to drive the LD to a level that clipping occurs with any regularity in optical digital cellular communications networks.

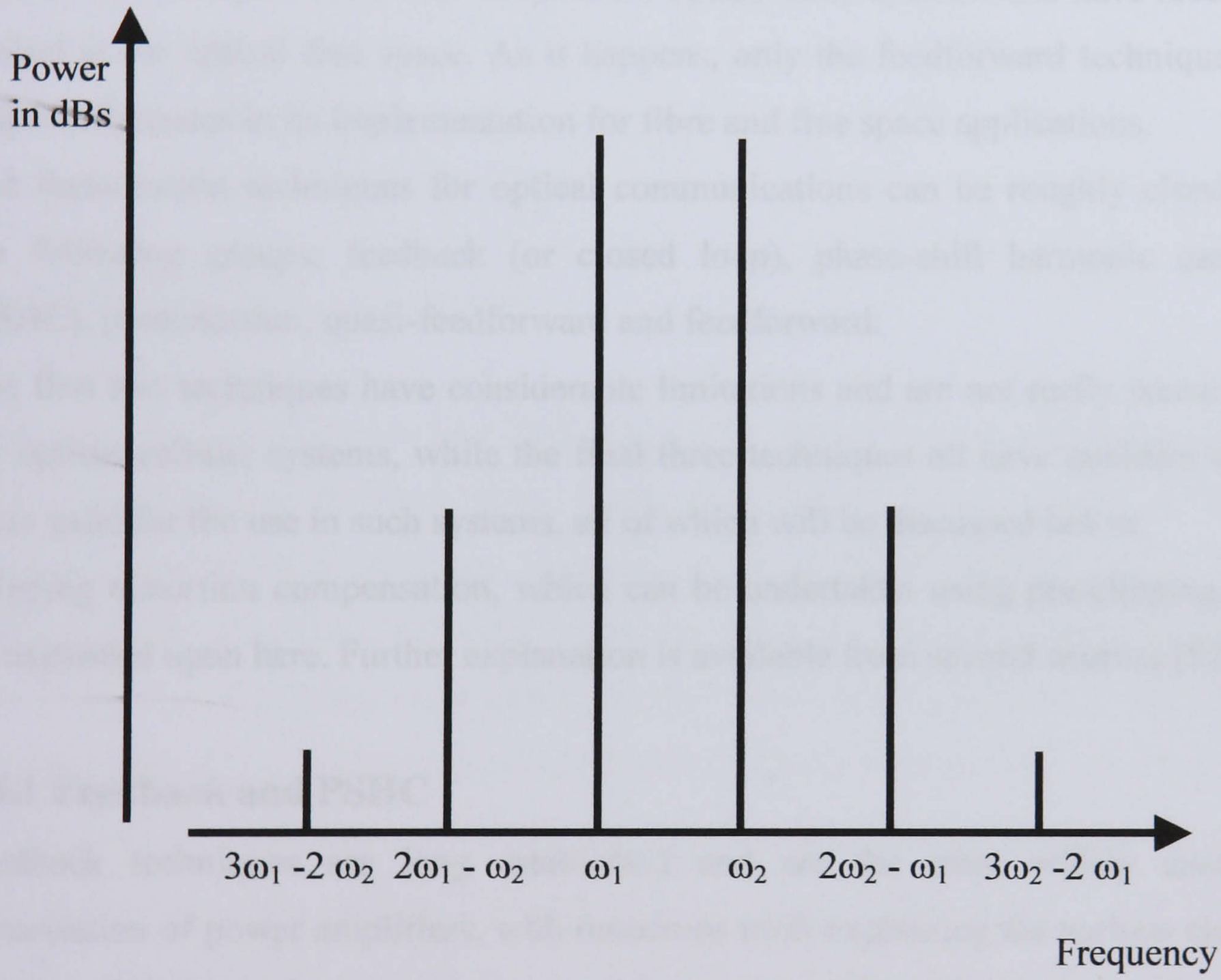


Fig 2.4 *Two tone intermodulation spectrum, with 3rd and 5th order terms*

2.5 Non-linear correction techniques

In this section some of the techniques that are available for the correction of non-linearities will be discussed, with the technique adopted for this study expanded upon in section 2.6.

Most of the non-linear correction techniques that have been established for optical communication were originally developed for microwave high power amplifier links. These suffer from many of the same sorts of non-linear distortion as lasers, though clearly the mechanisms involved in producing the distortion will differ. Thereafter, these correction techniques were then adapted for optical fibre systems and have recently been looked at for optical free space. As it happens, only the feedforward technique has any major differences in its implementation for fibre and free space applications.

The linearisation techniques for optical communications can be roughly classified into the following groups; feedback (or closed loop), phase-shift harmonic cancellation (PSHC), predistortion, quasi-feedforward and feedforward.

The first two techniques have considerable limitations and are not really serious options for optical cellular systems, while the final three techniques all have qualities that make them valid for the use in such systems, all of which will be discussed below.

Clipping distortion compensation, which can be undertaken using pre-clipping, will not be expanded upon here. Further explanation is available from several sources [52], [55].

2.5.1 Feedback and PSHC

Feedback techniques are long established and are the most widely used in the linearisation of power amplifiers, with numerous texts explaining the various closed loop designs, [56]. Feedback was also successfully used in early analogue video fibre systems. As Fig. 2.5 below shows, a proportion of the LD output is fed, via the photo receiver and inverter, back to the input of the system as a reference signal and combined with input signal. This combined signal now modulates the LD with the result of the noise and non-linearities of the LD being reduced. Both the noise level and the harmonics have shown to be reduced [57], with further reductions possible by the reduction of the delay in the feedback loop. However, in modern multichannel high frequency systems, the feedback

loop delay becomes serious restriction. At present, feedback systems can only offer a bandwidth of up to around 20MHz, which is clearly inadequate for the optical systems of today.

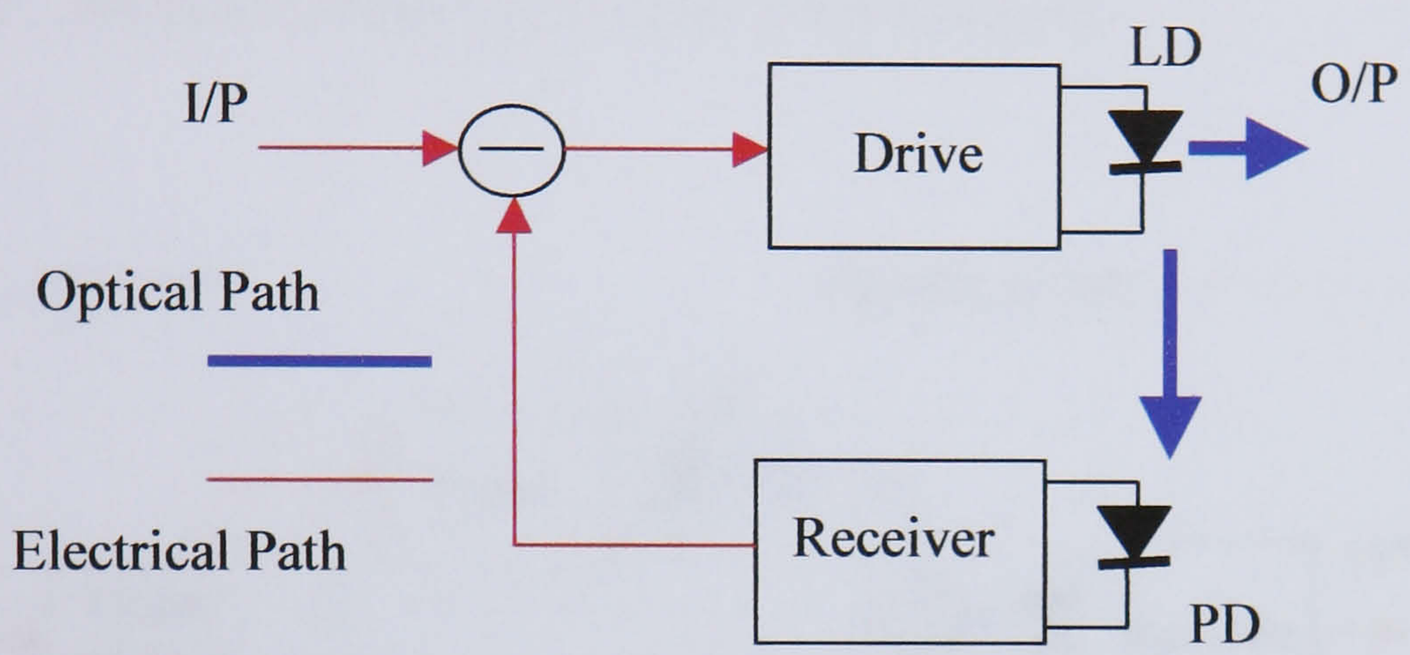


Fig. 2.5 *Optical Feedback technique*

Phase shift linearisation was, like feedback, used in the earliest of analogue fibre systems [58]. The technique involves using a pair of matching LDs to compensate for the distortion caused by a particular harmonic, such as the 2^{nd} or 3^{rd} order. As Fig.2.6 illustrates, the two LDs (and their respective drivers), are in parallel, with the input modulating signal arriving via a wideband, 3dB, 90^0 phase shifter that divides the signal into two parts. When the two signals are then combined, a phase shift exits between them. For the case of a $\pi/2$ phase shift in the fundamental, the result is a phase shift of π for the second harmonic and an amplitude of zero generated at the receiver. Meanwhile, the fundamental generated has a phase shift of $\pi/4$ and an amplitude enlarged by $\pi/2$. Other orders can be cancelled, with an adjustment to the phase shifter, with the *n*th order being cancelled with a phase angle difference of the two signals equal to π/n.

The first limitation with this technique is that only static non-linearities, based on the theory, appear to be cancelled. It will also not be suitable for multicarrier systems, due to intermodulation distortion not being compensated at the output frequencies which rely

upon differences of input frequencies. This is clearly inadequate for high frequency systems such as those like digital cellular.

The other main drawback is the need for two LDs with identical non-linear characteristics, which is very difficult to realise, and very expensive, as is the optical coupler. This limitation is only exaggerated when two different orders need cancelling, i.e. the 2nd and 3rd, because a second pair of LDs is then required.

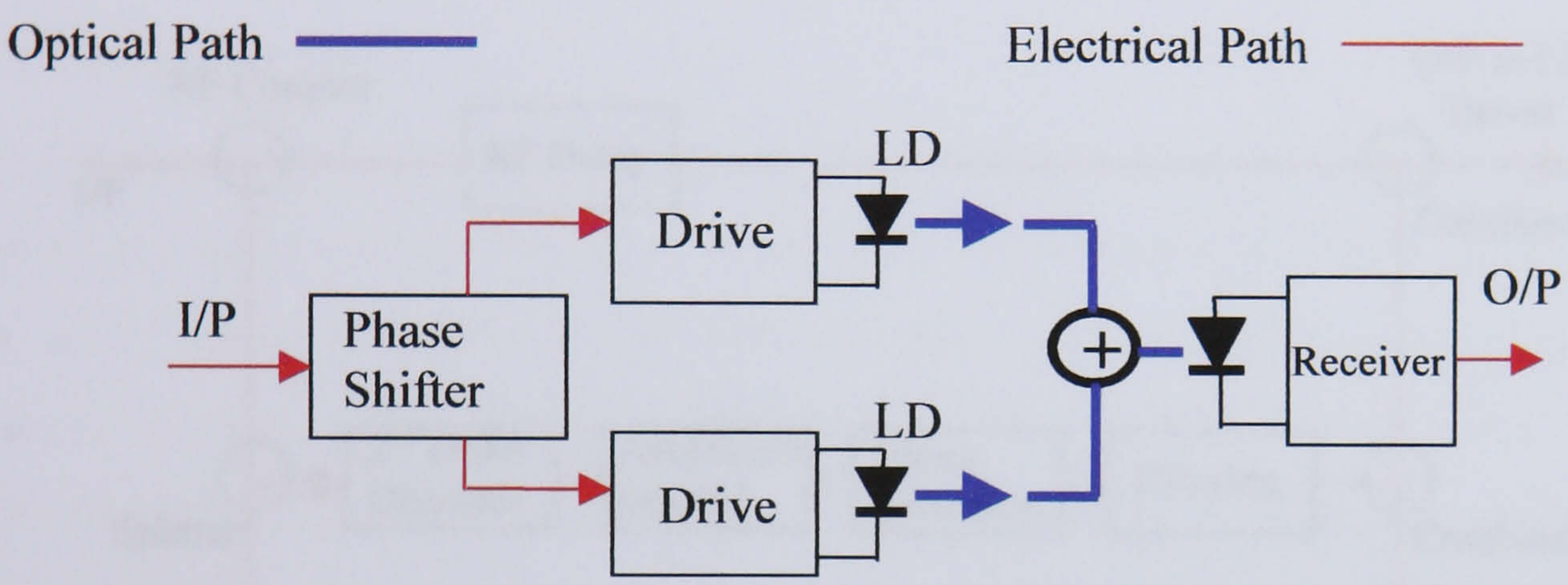


Fig. 2.6 *Phase Shift technique*

2.5.2 Predistortion

Of the three linearisation techniques listed above that are usable at the frequencies of optical digital cellular systems, the predistortion technique is generally regarded as the least expensive and complicated to implement. This is due to there being no extra optical components required, with all the linearisation being done electronically, which is not the case with feedforward and quasi-feedforward (as subsequent sections will illustrate). The predistortion technique has seen considerable research in recent years [42] and is already widely used in practical lightwave CATV systems [30]. Digital cellular systems have also seen some research with predistortion with the ROBIN project, (ROBIN - Radio Optical Broadband Integrated Network - was a collaboration between Bradford University, Orange PCS and Nortel Networks and was set up with the task of implementing a radio over fibre system for the DCS1800 2nd generation digital cellular network).

The idea of this technique is to electronically create a compensation signal using a pre-processor. This pre-processor consists of electronic circuitry that has a non-linear characteristic that is equal in amplitude but opposite in phase to that produced by the laser, with the effect of the two signals cancelling each other out. A predistortion system, cancelling for 2nd and 3rd order distortion is illustrated below in Fig. 2.7;

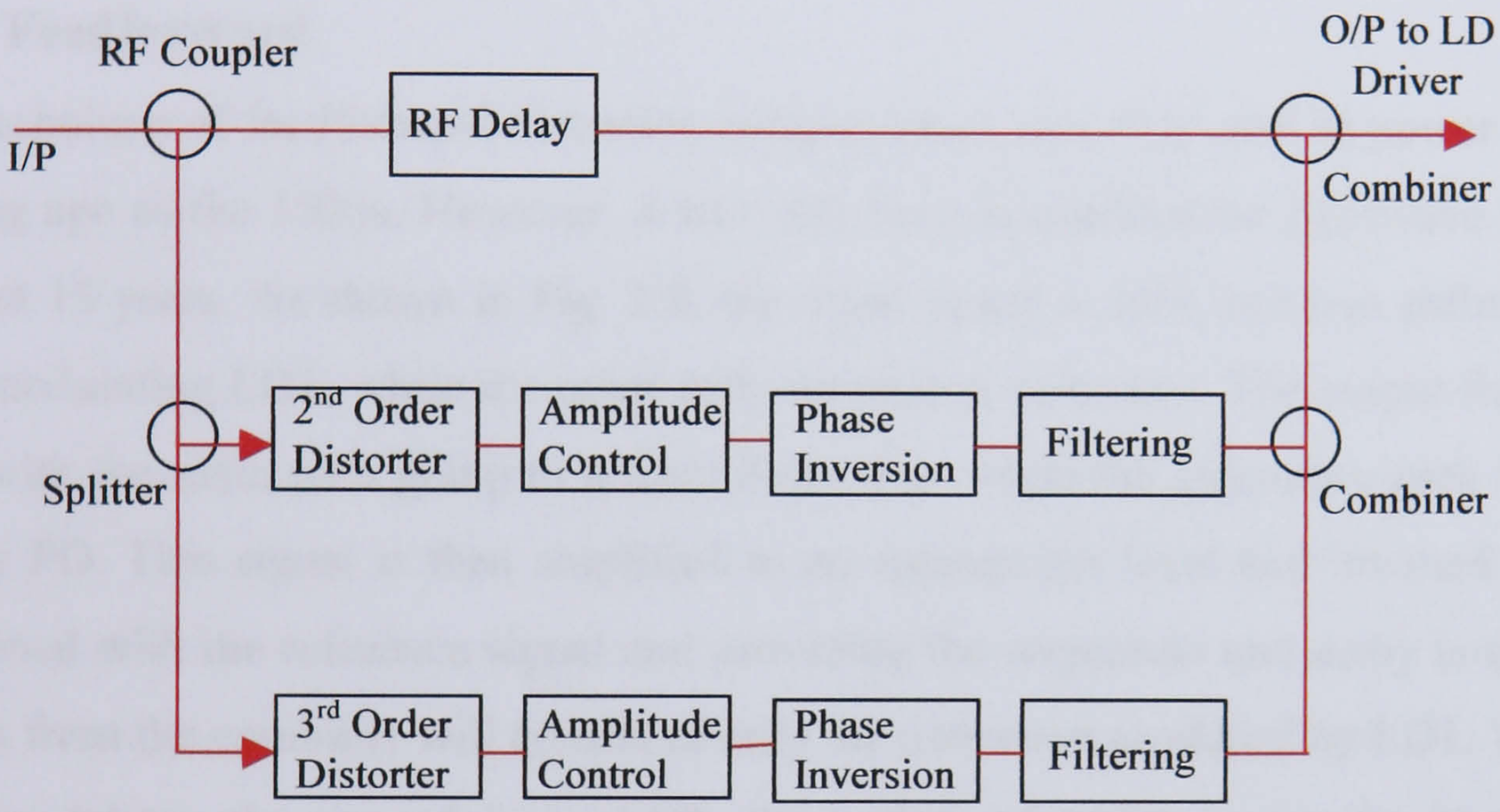


Fig. 2.7 *The Predistortion technique*

As can be seen above, the input signal enters a directional coupler to be separated into two paths, with the majority of the signal passing through to the delay unit. The remainder of the signal is then split again and fed into the circuits for the 2nd and 3rd order distorters. The signals are then appropriately adjusted for amplitude and phase and fed back into the main signal. Further details of this process can be found in several texts, [42], [52].

For this technique to work, a good understanding of the LDs characteristics is paramount, both theoretically, hence the use of the Volterra analysis, and practically, through experimental measurement.

The predistortion technique tends to yield a reduction of CSO and CTB of around 10dBs to 15dBs [42], [52]. However, this is difficult to maintain over a truly wide band of frequencies. Another drawback is the lack of ability of self-adjustment in the predistorter to match the variations of the LD operation with temperature and bias current, and those that occur with ageing. It is for these reasons, that, though inexpensive and comparatively simple, for actual performance, predistortion is considered inferior to the other broadband techniques.

2.5.3 Feedforward

The technique of feedforward distortion compensation was first used in power amplifiers as long ago as the 1920s. However, it has only been researched for lightwave systems in the last 15 years. As shown in Fig. 2.8, the input signal is split into two paths, with one path modulating LD1, while the other path becomes a reference. The output from LD1 is split with the main path going to a fibre delay line, while the secondary path is detected by the PD. This signal is then amplified to an appropriate level and inverted. It is then combined with the reference signal and providing the amplitude and delay is correct, the output from the combiner will consist of only the distortion produced by LD1. This signal then modulates the second laser, LD2, the output of which is mostly inverted LD1 distortion. The light from LD2 is then coupled into the delay fibre at the correct time such that, upon detection at the end of the link, the phasors representing the error signal and the LD distortion in the two optical signals cancel [47], leaving only the desired signal at the output of the receiver PD and amplifier circuit.

The first limitation with this technique is that the two LDs must be of different wavelengths to avoid coherent interference in the two signals. Consequently, the distortion compensation takes place in the receiving PD. This wavelength separation gives rise to chromatic dispersion in the fibre, resulting in a phase mismatch in the received signals and reducing the amount of distortion compensation [46]. To combat this, adjustments can be made to the electrical delay of the error signal for a given fibre dispersion. The non-linearities introduced by the second LD are the other main limitation to the level of compensation available. Though they are minor compared to those

produced by the main path LD, they do define the ultimate limitation to the system's performance, (though this could be improved by predistorting LD2).

The second laser also introduces additional RIN into the link, however, as the signal level is smaller than that of the main path LD, the RIN will also be smaller and therefore, not too significant. For FP lasers, a reduction of up to 10dB in RIN has been shown [59]. Meanwhile, reductions in CSO and CTB of around 15dB have been demonstrated over a frequency range of 100MHz to 1.5GHz [47].

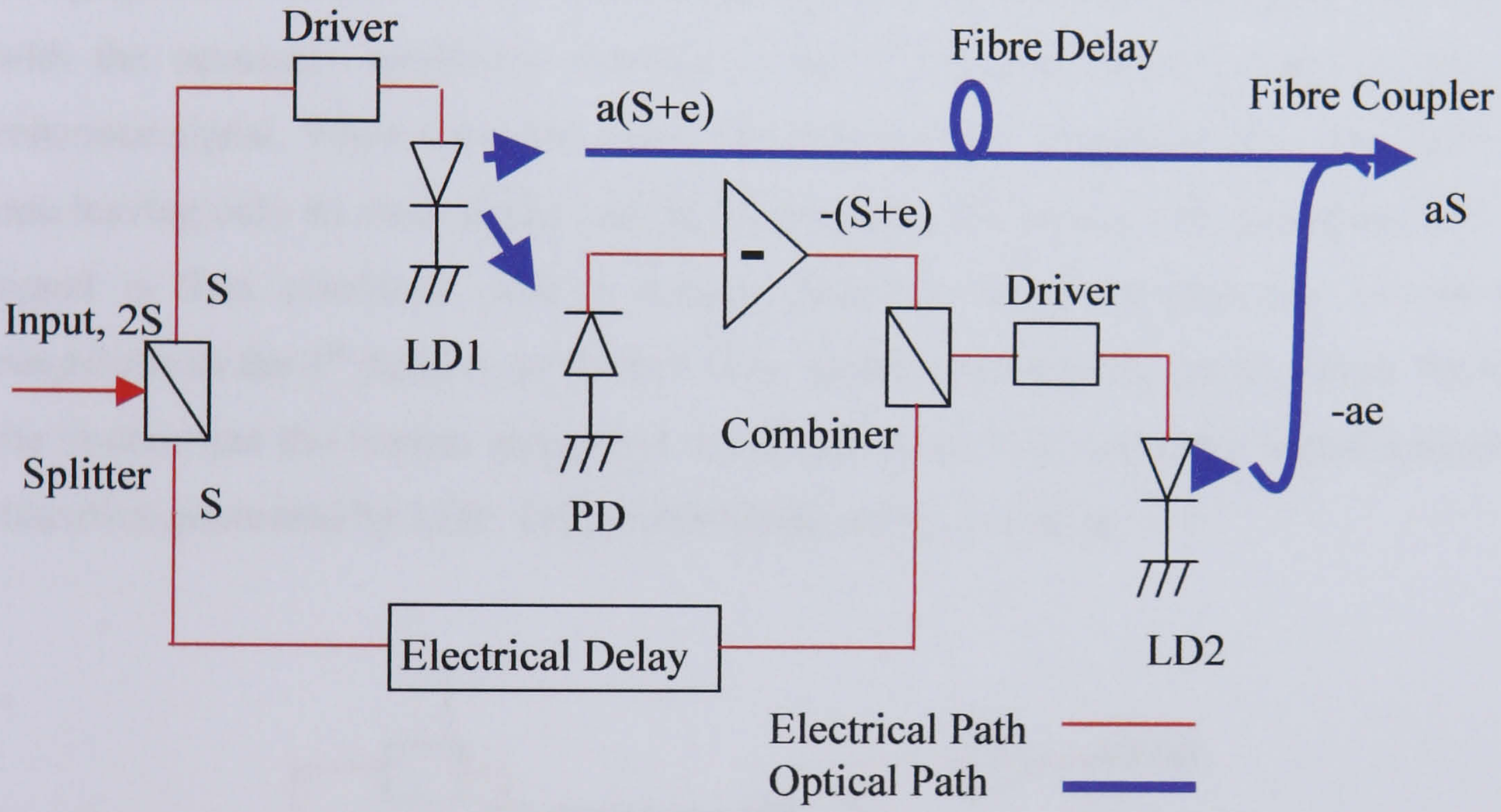


Fig. 2.8 *The Feedforward technique*

This significant broadband cancellation of non-linearities, as well as some cancellation of noise is the reason why overall, the feedforward technique is seen as by far the most promising and comprehensive technique for the cancellation of distortion in high capacity lightwave systems and will be the technique employed for this work. However, it is also far more complex and expensive than the other systems, due to the extra optical components, such as the coupler and fibre delay.

The system to be used in this study is of course free space, which will clearly not have many of the limitations of the above system, with there being no fibre. The changes that

will be required to the above system for free space operation will be expanded upon in section 2.6.

2.5.4 Quasi-feedforward

The quasi-feedforward technique is often seen as an amalgamation of predistortion and feedforward, having elements of both techniques. First proposed for a semiconductor light source (using LEDs) in 1977 [60], quasi-feedforward is theoretically the most effective of all the correction techniques discussed. However, it does have a major drawback, that will be alluded to in due course. The technique starts with the input signal being split three ways. The first path modulates an LD, that is then detected, inverted and with the necessary amplitude correction, mixed with the second path, an electrical reference signal. When these two are combined, the two fundamentals cancel each other out, leaving only an error signal with the inverted non-linearities introduced by LD1. This signal is then combined with a second reference signal (importantly of the same magnitude as the 1st path) to predistort LD2, which produces the output signal. Providing the system has the correct delay and amplitude correction, the error signal cancels the distortion generated by LD2. This is illustrated in Fig. 2.9 below.

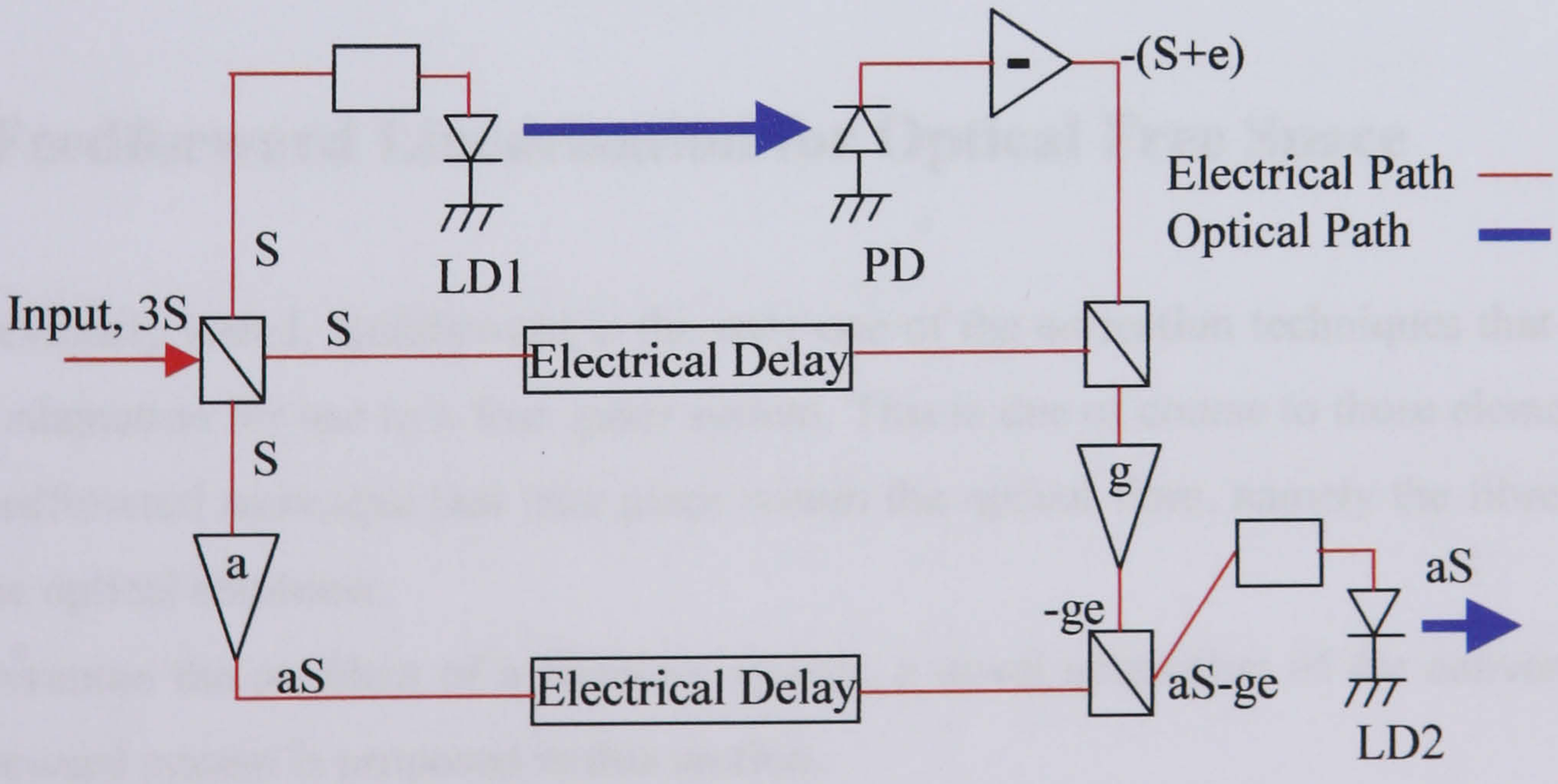


Fig. 2.9 *The Quasi-feedforward technique*

The advantages of quasi-feedforward over conventional feedforward are as follows. Firstly, as all the signal combination takes place pre-transmission, there is no need for the optical coupling, which as previously stated, can be both difficult and expensive. It also avoids the dispersion penalty associated with the use of two LDs with differing wavelengths. Lastly, the level of compensation is not limited by the distortion introduced by the second LD, as in feedforward.

One problem with quasi-feedforward is the additional RIN introduced by having two LDs in the system. The major drawback of this technique however, is that it depends on both the LDs having virtually identical non-linear characteristics for distortion compensation to take place. This is almost impossible to achieve (or at the very least, very expensive) as even lasers taken from the same batch wafer have large statistical variations in their characteristics. These variations can then be greatly aggravated by the effects of temperature fluctuations and laser ageing. This problem is then further increased when for a system to be used at higher frequencies, such as digital cellular, where the LD behaviour is more complex.

While improvements to LD fabrication may soon allow this technique to become fully realisable in the near future, the problem of finding matching LDs, makes it impractical for this project.

2.6 Feedforward Linearisation for Optical Free Space

As previously stated, feedforward is the only one of the correction techniques that needs some adaptation for use in a free space system. This is due of course to those elements of the feedforward technique that take place within the optical fibre, namely the fibre delay and the optical combiner.

To overcome the problem of a fibreless system, a novel adaptation of the conventional feedforward system is proposed in this section.

As can be seen if Fig. 2.10 below, the front end, or first loop, of the free space system is the same as for the fibre system (with the exception of a beam splitter being used to divide the free space beams). However, the second loop is vastly different. In the free

space system, both the main signal and error signal are transmitted and received separately. The two LDs will still have to be of slightly different wavelengths due to the fact that practically, they will be physically positioned quite close to each other and any coherent interference between the two signals still needs to be avoided.

This has the advantage of no chromatic dispersion (or at least so small an amount as to be insignificant) due to the optical signals travelling through free space and not fibre. This will make the timing less difficult to realise and is the first benefit to be seen of this technique.

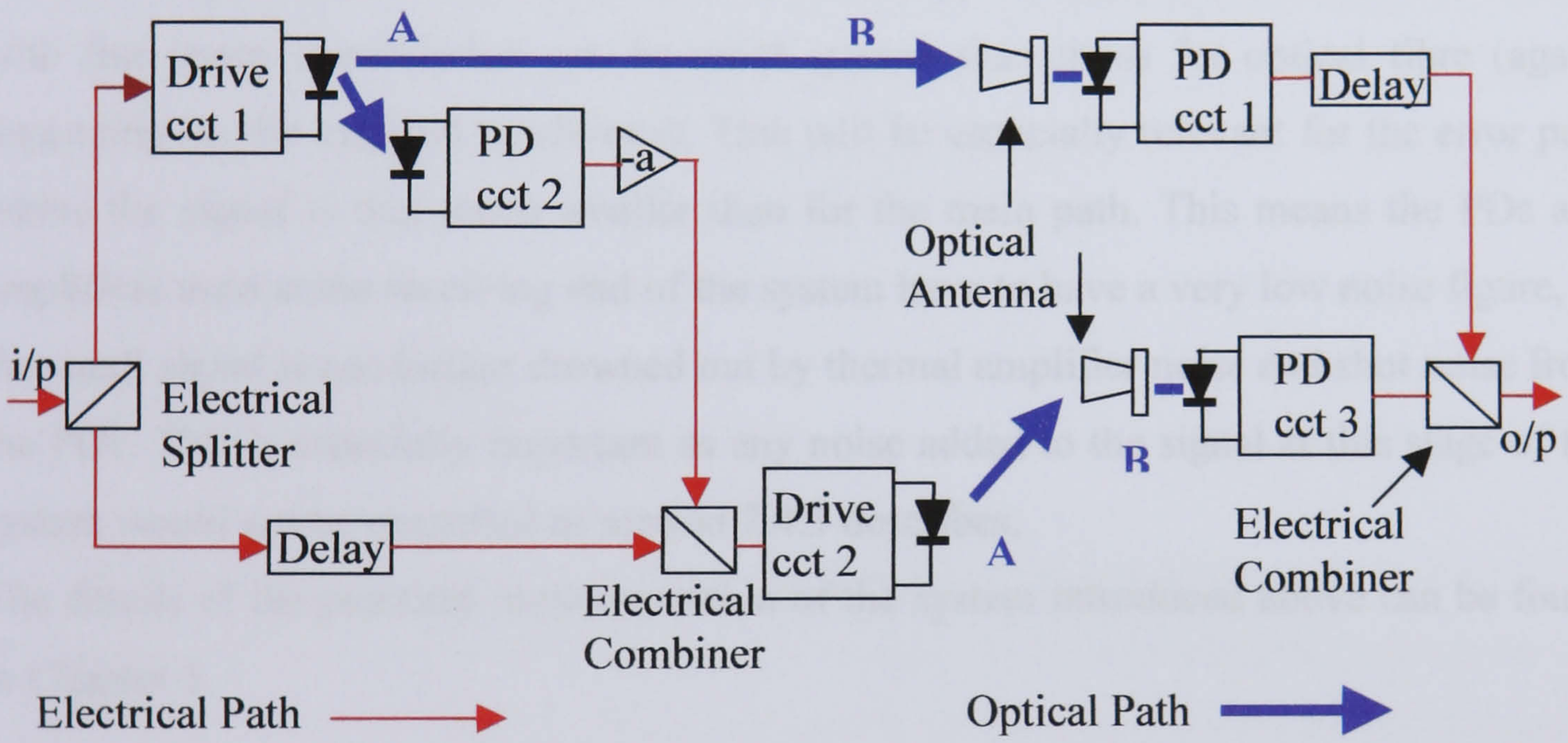


Fig. 2.10 *Optical Free Space Feedforward technique*

The two signals then pass through appropriate optical filters placed in front of their respective PD receiving units to further ensure against any coherent interference. The main path signal is then delayed for the required time before the two signals are electronically combined, with the opposite phase non-linearities (and some noise) cancelling each other out, leaving only the desired signal. Also included in this system is a pair of 'optical antennas', developed at the University of Warwick with patent pending, with one for each transmission path. The operation of these antennas will be further

expanded upon in Chapter 5, however, to summarise, they allow for the angle of incidence of an incoming light signal to be less than perfect, with all of the (non-reflected) signal transmitted, rather than some of it being refracted.

As can be seen from the above diagram, there is a need for three, rather than the standard two, receiving units. This will add to the cost of the system, however, compared to standard feedforward system, this cost is off set due to there being no need for the optical combiner and fibre delay, which are also expensive.

As stated above, some of the complexity associated with the feedforward technique is avoided due to the transmission being over free space rather than fibre, however the trade off will be an increased susceptibility to noise. As previously stated, the losses incurred with free space transmission can be much greater than those for optical fibre (again, depending on the external conditions). This will be especially relevant for the error path where the signal is that much smaller than for the main path. This means the PDs and amplifiers used at the receiving end of the system have to have a very low noise figure, so the small signal is not further drowned out by thermal amplifier noise and shot noise from the PDs. This is especially important as any noise added to the signal at this stage of the system would not be cancelled as section 2.4.3 describes.

The details of the practical implementation of the system introduced above can be found in Chapter 5.

In summary, this chapter has discussed the issues of system requirements, noise, non-linearities and the techniques available to solve them. Finally, a novel adaptation of the feedforward technique for use with a free space system has been introduced.

Chapter 3 – Taylor's Series, Perturbation Analysis and Volterra Series Expansions

In this chapter, the various techniques for modelling the optical free space feedforward linearisation system, introduced in the previous chapter, will be considered. The results of the subsequent simulations of the system for each technique will also be presented and discussed.

As stated in the previous chapter, the three techniques used to model the system will be the Taylor series, perturbation and Volterra series. The Taylor series is a relatively simple and widely used mathematical technique, whereas the perturbation and Volterra techniques both use expansion of LD rate equations, and hence are rather more involved.

3.1 The Taylor series

The Taylor series is one of a number of numerical progressions used frequently in mathematics [61]. Due to its simplicity, it is readily used within engineering, where it is often referred to as a 'power series', to model small signal non-linear behaviour in systems. Providing the system in question is memoryless, the Taylor series can give a very accurate description of non-linearities due to a potentially infinite number of terms [61]. However, in practise, as long as the non-linearities are not too extreme, only the first few terms of the series are needed, and it is rare to use any terms higher than the 5^{th} order. The Taylor series, up to and including the 5^{th} order (in the power series format), can be seen below;

$$3.v_o = a_1 v_i + a_2 v_i^2 + a_3 v_i^3 + a_4 v_i^4 + a_5 v_i^5 + \ldots. \quad \textbf{(3.1)}$$

where v_I and v_o are small, time varying quantities representing the RF input and output signals, and the a_n are scalar coefficients. For a two tone input case;

$$v_i(t) = v(\sin(\omega_1 t) + \sin(\omega_2 t)) \qquad \textbf{(3.2)}$$

In optical communications applications the LD is biased around a given point on the L-I curve (see Fig. 2.2). In this case the nonlinearity can be analysed by expanding the output power, Po(I), as a power series around the bias current point, I_{bias} [52]. Equation (3.1) thus becomes:

$$P_o(I) = a_1(I - I_{bias}) + a_2(I - I_{bias})^2 + a_3(I - I_{bias})^3 + a_4(I - I_{bias})^4 + \ldots. \qquad \textbf{(3.3)}$$

Where, for the two tone case:

$$I = I_{bias} + I_m(\sin(\omega_1 t) + \sin(\omega_2 t)) \qquad \textbf{(3.4)}$$

Where I_m is the modulation current amplitude. The I_{bias} term is not significant because of the the a.c. coupling of the system that effectively filters out any d.c. terms.

The system simulations have been undertaken in Matlab [62], with the different components of the system modelled mathematically and while the Matlab programs themselves are quite involved, the important components are expanded in the following.

The modelling of the laser diodes has conventionally been done using the Statz-de-Mars, or rate equations. These are two, coupled, differential equations that describe the interaction of electrons and photons within the active layer of the LD [34], [63].

Including any axial photons, (produced from spontaneous emission, which have little or no relevance once the LD is operating well above its threshold level), the rate equations for LDs are as follows:

$$\frac{dN}{dt} = \frac{I}{qV} - \frac{N}{\tau_n} - G(N,Q)Q \tag{3.5}$$

$$\frac{dQ}{dt} = \Gamma G(N,Q)Q - \frac{Q}{\tau_p} + \beta \tag{3.6}$$

where Q is the photon density, N is the electron density, G(N,Q) is the material gain (which is expanded upon in section 3.3), I is the current injected into the active region, V is the volume of the active region, q is the charge of an electron, τ_n is the recombination lifetime of the carriers (or lifetime of the excited state), τ_p is the photon lifetime, Γ is the optical confinement factor (given by the ratio of the active region and modal volume) and β is the spontaneous emission rate. The first term on the right of equation (3.5) that involves I, V and q, is also known as the 'pumping term' as it refers to electrons being pumped into the excited state.

If equations (3.5) and (3.6) are set to a steady state condition - i.e. the left hand sides set to zero - and solved for threshold conditions (where β term can be ignored due to the spontaneous emission being effectively negligible once the LD is in steady state), a general expression for the output power can be derived as:

$$P_o = (I - I_{th})Ep.\eta_D/q \tag{3.7}$$

where I is defined in equation (3.4), I_{th} is the threshold current, η_D is the differential quantum efficiency and Ep is the photonic energy as defined below;

$$Ep = hc/\lambda \tag{3.8}$$

Where λ = 850nm and of approximately 0.35, h is Plank's constant and c is the speed of light. Equation (3.7) essentially gives values for the scalar products in the power series and when combined with (3.3) gives an accurate expression for LD behaviour in the memoryless case. The threshold current I_{th} in (3.7), as well as I_{bias}, creates a d.c. offset. However, as the system is a.c. coupled, the term can be ignored.

As stated in Chapter 2, the photodiodes will be regarded as linear devices. The voltage out of a photodiode is as follows:

$$Vout = R(I_o + \Re P_{in}), \text{ and } \Re = q.\eta_p/Ep \qquad \textbf{(3.9)}$$

Where R is the PD load impedance of 50Ω, I_o is the dark response of the PD, which is a very small d.c. offset and hence can be ignored. P_{in} is the incoming signal power and is related to $P_o(I)$ subject to attenuation dependant on the distance travelled in free space (i.e. 20dB/km). $\Re$ is the responsivity of the PD and is made up of the electronic charge q, the internal quantum efficiency η_P (around 0.3) and the photonic energy Ep, as defined in equation (3.8). Together the $\Re$ and P_{in} create the input photocurrent and is a reverse of the electrical/optical conversion process of the LD.

The other components in the system, such as the electrical splitters and combiners, the coaxial cable, the beam splitter that splits the main beam and free space will be considered linear. In addition, any amplifiers used within the system will also be considered effectively linear. Hence, the amplifier gains were set to offset the losses incurred due to the other components within the simulation program, much as in the practical case.

The two-tone input for the simulations was set with frequencies at 2000MHz and 2050MHz, with a peak amplitude of 0.7V, giving a peak input current of 14mA, for a 50Ω LD load impedance. This gives a modulation index of 0.4 from equation (3.10) below:

$$M_{cur} = \frac{I_{peak}}{I_{bias} - I_{th}} \qquad \textbf{(3.10)}$$

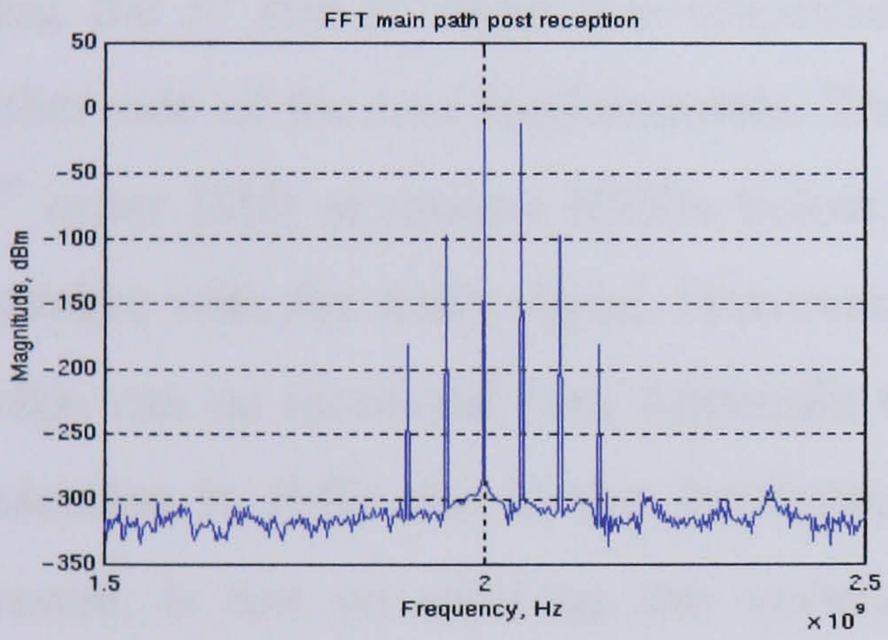
where m_{cur} is the modulation index, with I_{bias} set at 50mA and I_{th} set at 15mA. The higher the modulation index, the greater the performance of the link in the presence of noise. However, if the modulation index is set too high, the non-linearities generated will become more extreme. An overall index of around 0.2 to 0.4 (whether generated by one channel or the sum of many channels of smaller index) is generally considered to give a

good balance, and with this system needing as large a noise margin as possible, due to the comparatively low power signal, a value of 0.4 for the modulation index was chosen. The output of the main path LD is around 0dBm, or 1mW, which is close to the maximum allowed power of a laser output at this wavelength, i.e. around 800nm, to be laser eye safe for a collimated beam. This issue will be covered in greater depth in Chapter 5.

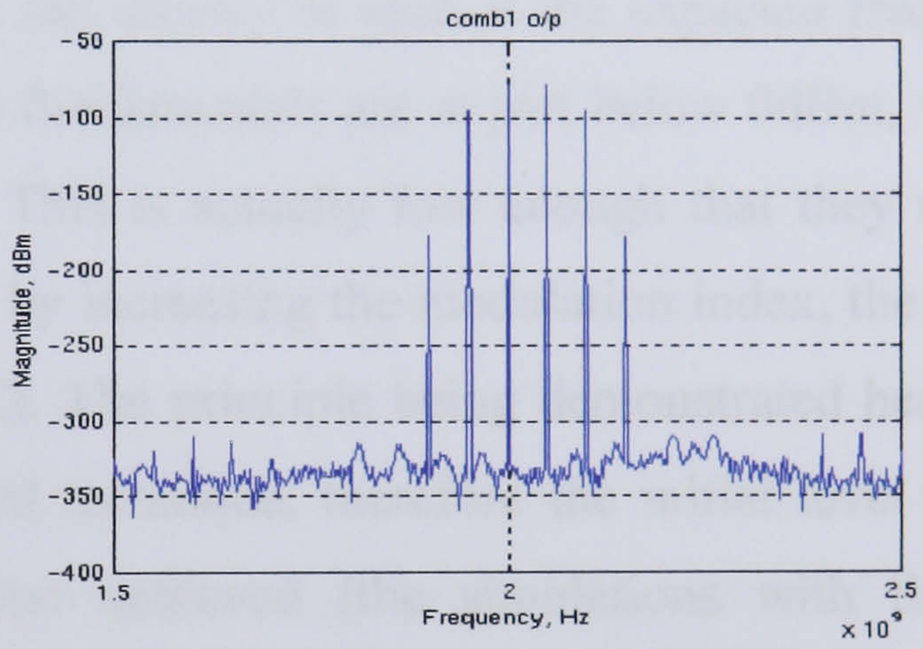
The simulations were run with both the 3rd and 5th order non-linearities included, using 8000 points, with FFTs taken at various stages within the system (as previously described in Figure 2.9) to display the required graphs.

Figure 3.1 below shows the results of the simulation at four different output sections of the system.

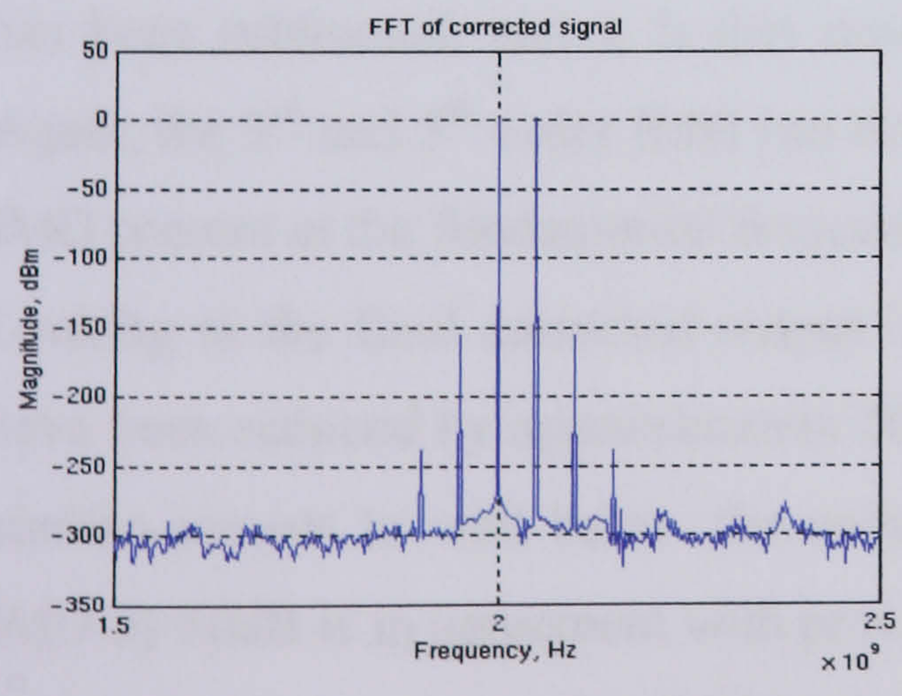
(a) Precorrective case.



(b) Non-linearity contribution.



(c) Final corrected output.



(d) IMD components only.

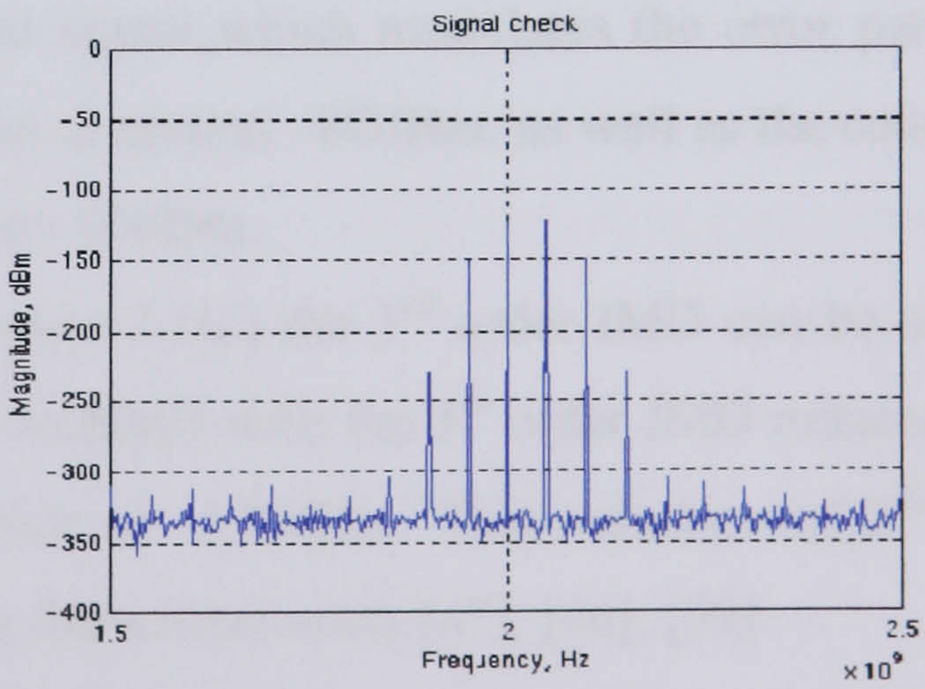


Fig. 3.1 *Results for Taylor's series optical feedforward non-linearity correction simulation*

The first point to make is that the above diagrams appear to show a noise floor of around –300dBm to –350dBm. From equations (2.8) and (2.10) in Chapter 2 it is known that the theoretical unit bandwidth noise floor at room temperature is equal to kT_o or –174dBm. Due to apparent limitations within Matlab, the noise properties of the system appeared too problematic to replicate faithfully. Although as previously stated in Chapter 2, the feedforward technique provides some cancellation of noise, the cancellation of IMD is very much the main focus of this work. Therefore the figures within this chapter that show simulation results will have an assumed noise floor of approximately –170dBm. The advantage of this scenario is that it allows the IMD behaviour below the noise floor to still be analysed.

Starting figure 3.1(a), with the output of the main receiving PD, i.e. the pre-correction case, the 3rd and 5th order non-linearities can clearly be seen at the expected frequencies either side of the two fundamentals. The fundamentals are at just below 0dBm, with the 3rd order IMD at around 80dBs below. This is actually low enough that they may not interfere with the main signal. However, by increasing the modulation index, the level of IMD_3 can be increased (see Appendix C). The principle being demonstrated here is the reduction in IMD due to the feedforward technique, therefore the initial level of IMD present is not so vital as the correction achieved (the simulations with the more sophisticated techniques show more realistic IMD amplitudes).

Figure 3.1(b) (top right) shows the signal due to non-linearities after the original signal has been subtracted, and it is this resultant signal which modulates the error path LD. Again, the 3rd and 5th order IMD can be seen at around –80dBm, as well as the combined IMD present at the fundamental frequency at –60dBm.

Looking at the final corrected output in figure 3.1(c) the 3rd order IMD can be seen to have been reduced by approximately 50dB to 60dB with the 5th order IMD reduced by a similar amount to well below the noise floor of –170dBm. This reduction of 3rd order IMD by 50dB is in agreement with previous theoretical work [47], [46], [59].

Figure 3.1(d) is the result of taking away a theoretically perfect signal away from the final corrected output, to leave just the present IMD, and again shows an approximate reduction of 50dB plus when compared to the output of the combiner above. This graph

is included as an additional demonstration of the feedforward technique working. However, this would not be realisable in a practical system.

From the figures in 3.1 above, it can be concluded that for the Taylor series model, the optical feedforward approach does theoretically reduce the IMD_3 present by a significant amount as indicated by theoretical simulation.

As is common in the simulation of LD communications systems, the above simulation can be repeated with different input scenarios. The most common parameters to alter are the bias current and the modulation index, via the modulation current. However, as shown in equation (3.10), the two are intrinsically linked. Therefore it was decided that, due to the ease of implementation within the Matlab program, I_{peak} in (3.10) would be adjusted to give a different modulation index. These simulations, (the graphs of which can be seen in Appendix C [figures C1 and C2]), show the results of where $m_{cur} = 0.2$ and 0.6 respectively. For a value of $m_{cur} = 0.2$, figure C1 shows a slight drop in the fundamental signal but a larger drop in the IMD_3 of around 20dB, when compared to the $m_{cur} = 0.4$ case (with the IMD_5 now below the noise floor). This is much as expected. However, the important point to note is that the IMD cancellation due to the feedforward technique is virtually unaltered at just over 60dB. For the case of $m_{cur} = 0.6$, the pattern is reversed, with, again, little change in the level of the fundamental, but with an increase in IMD products. However the system still corrects the IMD for a reduction of just over 50dB.

The slight decrease in reduction of IMD is not too surprising, as a modulation index of 0.6 could be considered as 'over modulation', and the resulting intermodulation distortion could be large. If the distortion becomes too large i.e. to the point of the system being strongly non-linear, the small signal analysis used could start to break down. In this case, the higher powers of non-linearity need to be taken into account and other analysis techniques such as the harmonic balance and time-domain approaches such as those shown in the subsequent sections, may be required.

Before moving on to the more sophisticated techniques, the above $m_{cur} = 0.4$ simulation was repeated, this time with the 2^{nd} and 4^{th} order distortion included. The viewing

frequency range was thus extended to include the extra non-linearities. The reasoning for this was to see if the optical feedforward technique would correct for these harmonic distortions, if required, the results being shown in figure 3.2 below.

Comparing the pre-correction and post correction results above, it can clearly be seen that the 2nd order distortion at around 4GHz is equally well, if not better, (theoretically) compensated for as that for the 3rd order distortion. This again agrees with previous findings with the feedforward technique in optical fibre systems [47], [46], [59].

(a) Precorrective case.

(b) Final corrected output.

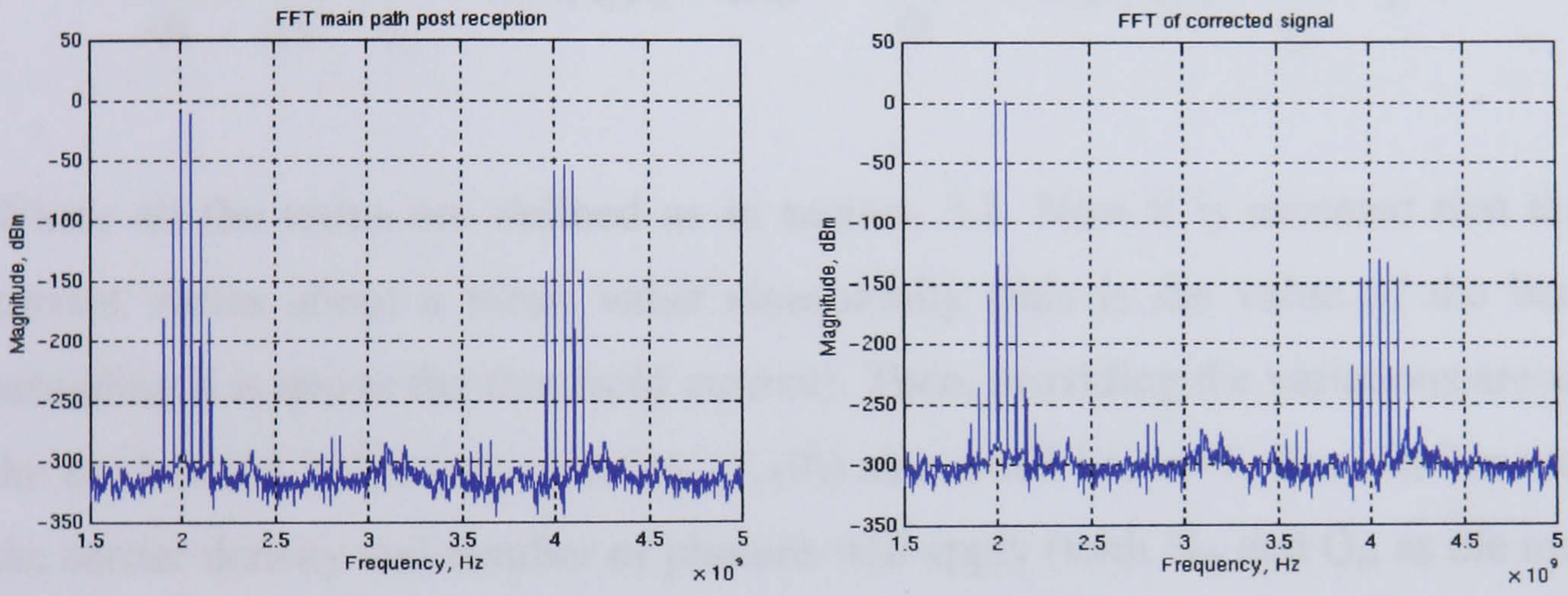


Fig. 3.2 *Expanded frequency scale Taylor's series simulation results*

3.2 Perturbation analysis

The term perturbation is defined as 'the action of altering or moving the normal or regular state of a system or process'. Hence using the rate equations above as a starting point, perturbation analysis adapts them as so to provide a reasonable representation of dynamic distortion, whilst avoiding the mathematical complexity that can be associated with a more all encompassing numerical approach [48].

The technique of perturbation analysis for laser diode non-linearities first came to the interest of engineers in the mid 1980s for the modelling of sub-carrier multiplexed

systems (SCM) [31], [64], with further subsequent development by Helms [65] in the early 1990s. Since then, however, the Volterra series would appear to have proved a more popular technique for the analysis of LDs. Nonetheless perturbation is still a valid technique, and comparisons between techniques are always of interest.

A detailed expansion of the derivation can be found in Appendix A, with the following serving as a summary. The technique begins with the LD rate equations of (3.5) and (3.6) shown below:

$$\frac{dN}{dt} = \frac{I}{qV} - \frac{N}{\tau_n} - G(N,Q)Q \quad \text{and} \quad \frac{dQ}{dt} = \Gamma G(N,Q)Q - \frac{Q}{\tau_p} + \beta$$

Where all the terms are defined as in section 3.1. Now it is assumed that the injected current varies about a mean value sinusoidally (this is the value of the bias current, providing it is above the threshold current). Then, providing the variations are small, with the modulation index (see equation (3.10)) also small i.e. m<1, then similar variations in the carrier density and number of photons will apply (with N_{th} and Q_m as the mean values respectively).

The expressions for I(t), N(t) and Q(t) for a LD with a sinusoidal current input are as follows:

$$I(t) = I_{bias} + Re(\Delta I e^{j\omega t}) \tag{3.11}$$

$$N(t) = N_{th} + Re(\Delta N e^{j\omega t}) \tag{3.12}$$

$$Q(t) = Q_m + Re(\Delta Q e^{j\omega t}) \tag{3.13}$$

These expressions can then be fed back into the rate equations of (3.5) and (3.6). By neglecting any terms higher than the 3rd order and the noise sources, a set of linearised

equations in the frequency domain can be found for N and Q in terms of the input frequency ω. The higher order distortion products can then be derived by using these fundamental solutions as the driving terms in the linearised equations. With the laser's fundamental normalised transfer function H(jω) defined as [65]:

$$H(j\omega) = \frac{1}{(j\omega/\omega_r)^2 + j\omega/j\omega_d + 1} \tag{3.14}$$

where ω_r is the relaxation resonance frequency and ω_d is the damping frequency (see Appendix A), the expressions for the intermodulation products ID2 and ID3 can then be found at the frequencies $\omega_{ID2} = \omega_1 \pm \omega_2$ and $\omega_{ID3} = \omega_1 \pm \omega_2 \pm \omega_3$ respectively as:

$$ID2 = \frac{1}{2} \left| m_{opt} H(j\omega_{ID2}) A(j\omega_{ID2}) \right| \tag{3.15}$$

$$ID3 \text{ (or } IMD_3) = \frac{1}{4\omega_r^2} \left| m_{opt}^2 (j\omega_{ID3} + 1/\tau_n) H(j\omega_{ID3}) C(j\omega') \right| \tag{3.16}$$

where m_{opt} is these optical modulation index. Definitions of the functions $A(j\omega_{ID2})$ and $C(j\omega')$ can be found in Appendix A, where an expression for m_{opt} can also be found. However, in general, m_{opt} is considered to be approximately the same as the input current modulation index, m_{cur}.

In this study it is only the 3rd order intermodulation products that are of major concern. The above ID3 expression is simply adapted so $\omega_1 = \omega_3$. Hence, for IMD_3, $\omega_{IMD} = 2\omega_1 \pm \omega_2$. The ID2 expression is included for completeness.

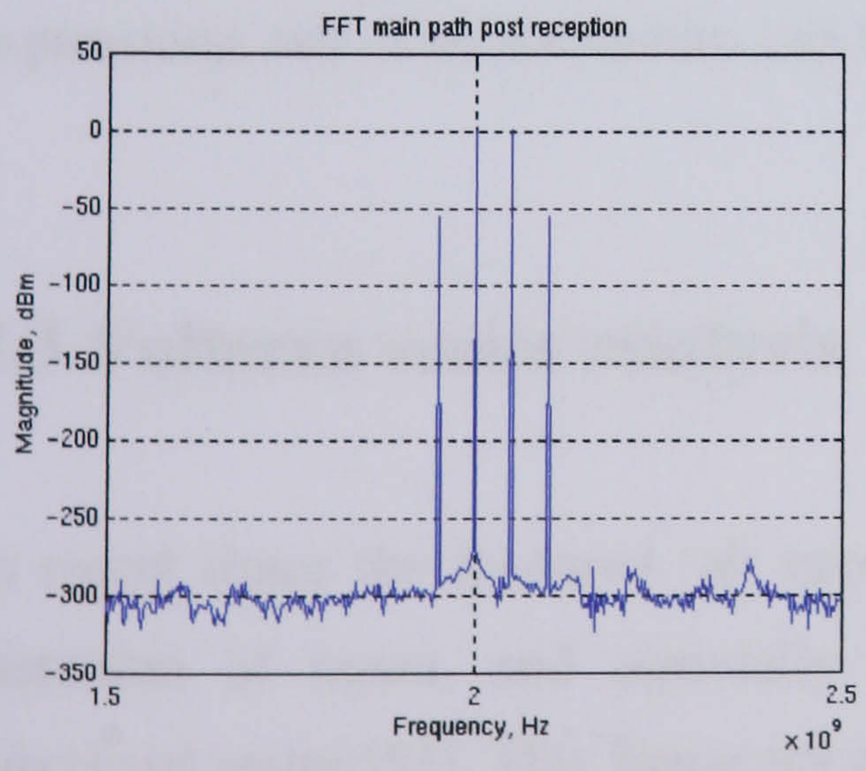
The expressions for the fundamental and the IMD_3 are now incorporated into the Matlab program as for the Taylor's series, with the same scalar expressions seen in equations (3.7) and (3.9) used. The rest of the system set up is the same - only the LD transfer function has changed - with the perturbation expression for the fundamental plus IMD, replacing the Taylor's series version.

Unlike the case for the Taylor's series, where the 5th order intermodulation product was included, only the 3rd order IMD is shown. The results of the pre-correction system and post correction system can be seen in Figure 3.3 below.

As stated above, the input values are the same as for the Taylor's series. However, values for some extra terms do need to be defined. The value of m_{opt} in the IMD_3 transfer function was set at 0.4, τ_n at 0.68×10^{-9}s, τ_p at 0.59×10^{-12}s and the damping rate, γ is at a value of 0.03 (see Appendix A).

When compared with the results for the Taylor's series, it can be seen the IMD_3 is significantly higher, approximately 50dB below the level of the fundamental. However, importantly, modelling the optical feedforward non-linearity correction system using perturbation analysis shows a reduction in IMD_3 of around 50 to 60dBs. This, again, agrees with the simulation results for the Taylor's series. The presence of very small (–250dB) higher order non-linearities is likely to be due to the non-linear characteristics of the correction path laser, which, as previously stated in Chapter 2, is one of the ultimate limitations of the feedforward correction approach.

(a) Precorrective case.



(b) Final corrected output.

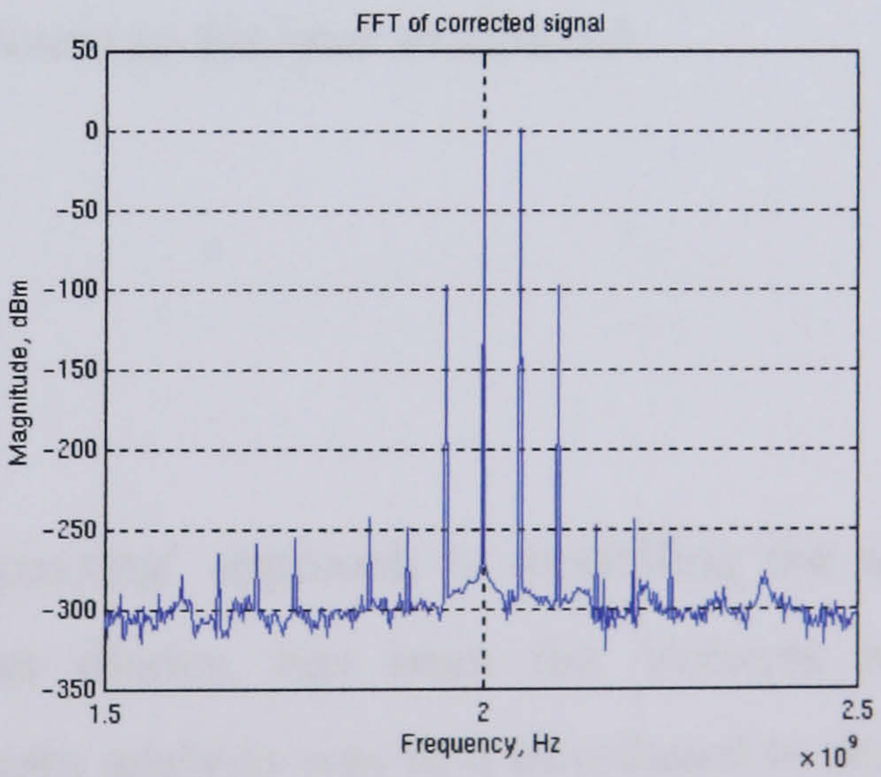


Fig. 3.3 *Results for perturbation analysis of the optical feedforward non-linearity correction simulation*

The simulation was again repeated for $m_{cur} = 0.2$ and 0.6, with the m_{opt} value also set at 0.2 and 0.6 correspondingly. Using the results shown in Appendix C (figures C3 and C4), it can been seen that, for $m_{cur} = 0.2$ the same pattern as for the Taylor's series results is followed, with little change in the fundamental, but with a lower level of IMD_3. However. the reduction in IMD_3 is slightly more pronounced, being around 60dB.

When the modulation indexes are set at 0.6, the initial IMD_3 is again increased, and is only around 40dB below the fundamental. The amount of reduction of IMD_3 seen. post correction, is also reduced to around 40dB, which is still very significant. However, the higher order distortion noticed for the m_{cur} (or m_{opt}) = 0.6 case, has increased to approximately –200dB. While this level is still not above the noise floor, it does suggest that too great a modulation depth may indeed cause problems. The reduction in IMD_3 cancellation is most likely to be due to the limitation in feedforward correction of the IMD_3 introduced by the correction path LD. However, it may also possibly be due to the small signal assumptions starting to break down when over modulation is used.

It is worth noting at this point that, within the Matlab perturbation analysis program, both the current and optical modulation indexes are included, which may exaggerate their effect. Also to enable the perturbation analysis for these particular simulations on the optical feedforward non-linear correction system to be undertaken, some values had to be calculated by expressions derived in the Volterra analysis. Further discussion of these expressions, and of all the results, can be found in the later section 3.5.

3.3 Volterra series analysis

In recent times the favoured 'all encompassing' approach to modelling the non-linear distortion of lasers, and especially laser diodes, has been the Volterra non-linear functional series [53]. This frequency domain analysis was first introduced to the analysis of non-linear circuits in 1942 by Wiener [66], and, since then, the theory has been expanded and applied in a general way to a number of problems in non-linear systems. It is particularly useful in calculating small (but troublesome) distortions produced in various types of communication system [67], [68]. Details of the Volterra analysis

method can be found in numerous publications [67], [68], [69], [70] with its use in a feedforward linearisation system for fibre undertaken by Hassin and Vahldieck in 1993 [46]. However, a full explanation of the mathematical theory of the Volterra method, and its application to LD distortion, can be found in Liang [52].

In the following a brief outline of the more important expressions is given with further, more detailed, analysis in Appendix B1.

As stated previously, in a non-linear LD system in which the distortion is dependent upon the frequency of the modulation signal, i.e. for those frequencies above approximately 100MHz, the non-linearity can be best represented by the Volterra functional series, also known as a power series with memory.

Beginning with the premise that the system may be analysed as a 'black box' system, the input is the function x(t) and has a corresponding output function y(t). The two can be related by a series that was first studied by the mathematician Vito Volterra [71], hence the name of the technique. By a sequence of Fourier transforms and inverse Fourier transforms, with further substitution and manipulation, an expression for y(t) in terms of x(t) via the transfer functions H(f) can be found. This is summarised in figure 3.4 below.

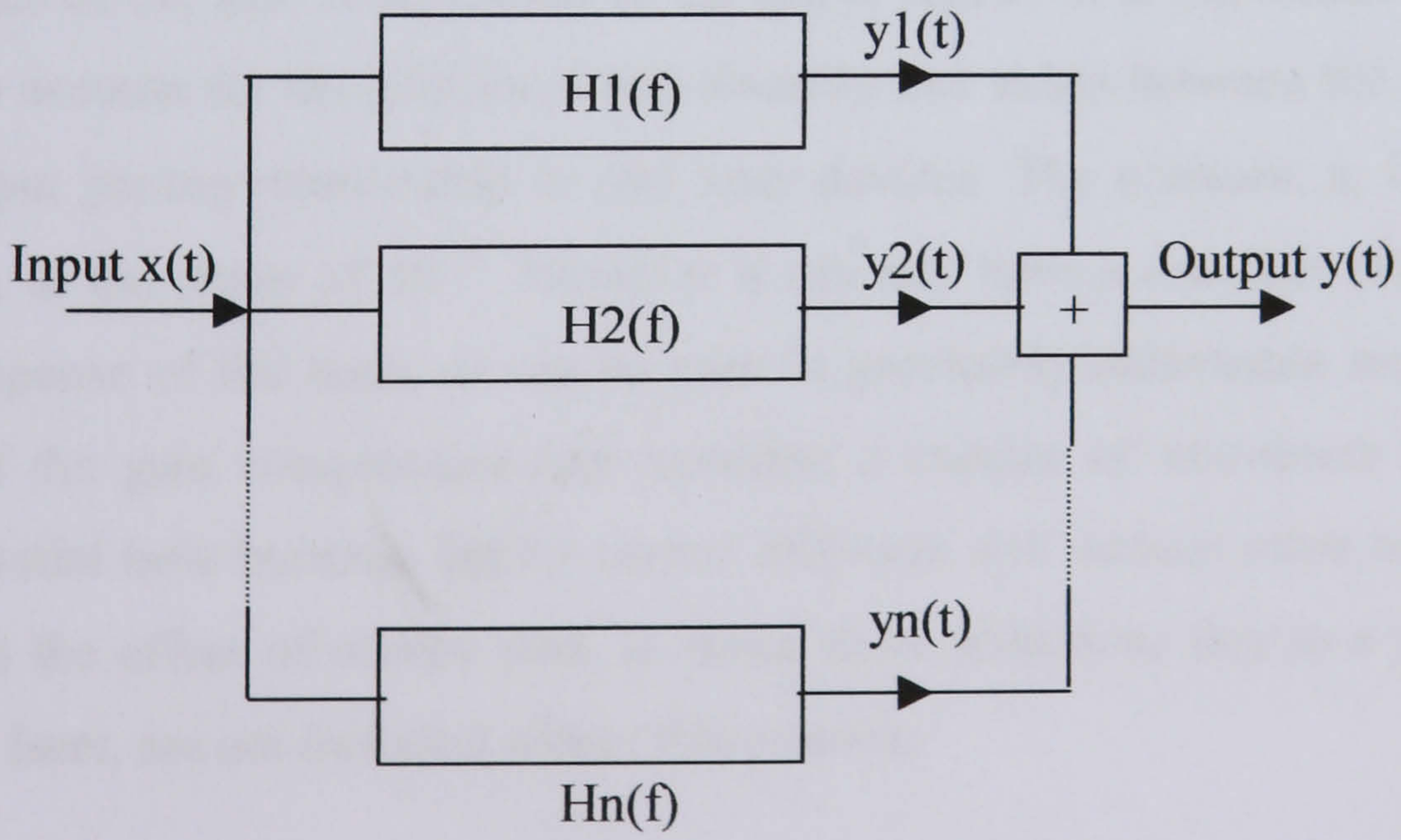


Fig. 3.4 *Representation of a weak non-linear system characterised by a Volterra series*

Now, starting with the LD rate equations of (3.5) and (3.6) as shown below:

$$\frac{dN}{dt} = \frac{I}{qV} - \frac{N}{\tau_n} - G(N,Q)Q \quad \text{and} \quad \frac{dQ}{dt} = \Gamma G(N,Q)Q - \frac{Q}{\tau_p} + \beta$$

where as before, Q is the photon density, N is the electron density, I is the current injected into the active region, V is the volume of the active region, q is the charge of an electron, τ_n is the recombination lifetime of the carriers (or lifetime of the excited state), τ_p is the photon lifetime, Γ is the optical confinement factor, β is the spontaneous emission rate. However, this time expanding upon the term for material gain gives $G(N,Q) = g'(N-N_0)$, where g' is the differential gain and N_0 is the transparency carrier density. The differential gain can be further expanded to $g' = g(1-\varepsilon Q)$, giving the overall expression as:

$$G(N,Q) = g(1-\varepsilon Q)(N-N_0) \qquad \textbf{(3.17)}$$

Where g is the optical gain factor, which is determined by material factors. The term ε is the gain suppression parameter and is a phenomenological constant which identifies the characteristics of the gain compression of the active region. It is introduced into the rate equations to account for the additional non-linearity that exists between the input current and the output photon relationship in real laser devices. The constant, ε, is generally a small value, in the realm of 10^{-23}. However it can still have a dramatic effect upon the dynamic response of the laser, as can be seen in previously undertaken work [52]. The inclusion of the gain compression can represent a number of non-linear mechanisms, including spatial hole burning, lateral carrier diffusion and various other non-linearities [72], though the effect of others, such as those from reflections due to a pigtail or the laser mirror facet, are not included within this constant.

As previously stated, the two LD rate equations describe the non-linear relationship between the injected carriers and the photons in the laser cavity. With the laser biasing point well above the threshold condition, the spontaneous process can be effectively

ignored with respect to the stimulation process i.e. giving $\beta = 0$, the two rate equations of (3.5) and (3.6) being combined to form a third equation that expresses the injection current I in terms of the photon density as can be seen below:

$$I - I_{th} = \frac{V'}{\Gamma}\left(\frac{dQ}{dt} + \frac{Q}{\tau_p}\right) + \frac{V'}{\Gamma g}\frac{d}{dt}\left[\frac{\frac{dQ}{dt} + \frac{Q}{\tau_p}}{(1-\varepsilon Q)Q}\right] \qquad \textbf{(3.18)}$$

where $V' = qV$ and where it has been given that $I_{th} = qVN_0/\tau_n$. Due to the relatively large value of τ_n, a $1/\tau_n$ term has been set to zero for simplicity. Further explanation of this can be found in Appendix B1. Equation (3.18) gives the somewhat unusual description of the input, I, as a function of its output, Q, with the steady state solution showing a linear relationship of the dc parts. For the analysis of the non-linear model, the time varying part of (3.18) needs to be examined. Firstly, both the injection current and photon density can be separated into ac and dc components:

$$I(t) = I_0 + i(t) \quad \text{and} \quad Q(t) = Q_0 + q(t) \qquad \textbf{(3.19)}$$

where I_0 is the laser biasing current, i(t) is the modulation current, Q_0 is the steady state photon density associated with I_0, and q(t) is the time varying photon density associated with i(t). Next, the denominator of the last term in (3.18) is substituted by its Taylor's series expansion around the steady state, (as shown in Appendix B1). When this expansion, along with the expressions of equation (3.19) are substituted into (3.18), the resultant expression is:

$$\begin{aligned} i(t) = {} & A + Dq(t) + Eq'(t) + Fq''(t) + Mq(t)q'(t) + Nq(t)q''(t) + Nq'^2(t) \\ & + Sq^2(t)q'(t) + 2Gq(t)q'^2(t) + Gq^2(t)q''(t) + \ldots. \end{aligned} \qquad \textbf{(3.20)}$$

where A, D, E, F, M, N, S and G are expressed in the terms of the laser parameters and Q_0, all of which are explained in detail in Appendix B1.

The Volterra series can now be used to relate i(t) and q(t) in the non-linear distortion modelling, as follows [69]:

$$i(t) = \sum_{n=1}^{\infty} \frac{1}{n!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g_n(u_1 \ldots u_n) \cdot \prod_{k=1}^{n} q(t - u_k) du_1 \ldots du_n \qquad \textbf{(3.21)}$$

in which, $g_n(u_1 \ldots u_n)$ is called the *n*th order Volterra kernels of the system. Equation (3.21) gives the true input current in terms of the true output, and is called the inverse laser system. To acquire the forward Volterra transfer functions, $H_n(\omega_1, \ldots \omega_n)$, which treat i(t) as the input of the system, the inverse Volterra transfer functions, $G_n(\omega_1, \ldots \omega_n)$, for the system which treat q(t) as their inputs, have to be derived first [52].

Using the harmonic input method [73], the first three G_n terms (which, assuming small signal analysis, are all that is required) are calculated from equation (3.20) (see Appendix B1) resulting in the first, second and third order Volterra transfer functions, as seen below:

$$G_1(\omega) = D - F\omega^2 + j\omega E \qquad \textbf{(3.22)}$$

$$G_2(\omega_1, \omega_2) = j(\omega_1 + \omega_2)M - (\omega_1 + \omega_2)^2 N \qquad \textbf{(3.23)}$$

$$G_3(\omega_1, \omega_2, \omega_3) = j2S(\omega_1 + \omega_2 + \omega_3) - 2G(\omega_1 + \omega_2 + \omega_3)^2 \qquad \textbf{(3.24)}$$

From the above, using the *n*th order inverse [71] and the harmonic balance method, the forward Volterra transfer functions required for actual system modelling, $H_n(\omega_1, \ldots \omega_n)$, can be calculated from G_n. The first order functions have a simple relationship, as in linear systems:

$$H_1(\omega) = \frac{1}{G_1(\omega)} = \frac{1}{D + j\omega E - F\omega^2} \qquad \textbf{(3.25)}$$

The expressions for the second and third order functions $H_2(\omega_1,\omega_2)$ and $H_3(\omega_1,\omega_2,\omega_3)$ are calculated in a similar fashion. However, as these rather more involved, they are not listed here and can be seen in Appendix B1.

If the output properties of the Volterra systems are now driven by a two tone harmonic input of $x(t) = P\cos(\omega_1 t) + Q\cos(\omega_2 t)$, where $\omega_{1,2} = 2\pi f_{1,2}$, it can be shown [52] that the fundamental, the second order intermodulation distortion of the frequency $(\omega_1 + \omega_2)$ and the third order intermodulation distortion of the frequency $(2\omega_1 - \omega_2)$, detected at the photo receiver, are as follows:

$$\text{Fundamental} = PH_1(\omega_1)/2 \quad \textbf{(3.26)}$$

$$IMD_2 = P^2H_2(\omega_1,\omega_2)/4 \quad \textbf{(3.27)}$$

$$IMD_3 = P^3H_3(\omega_1, \omega_1, -\omega_2)/16 \quad \textbf{(3.28)}$$

where P is the optical power, as calculated in equation (3.7). As with the perturbation technique in section 3.2 the definition for IMD_2 is included for the purposes of completeness. However, unlike the perturbation technique, the Volterra analysis provides separate expressions for the 2^{nd} and 3^{rd} order harmonic distortion, as well as for intermodulation distortion. The above expressions are normalised with respect to $H_1(0)$ and are written in a Matlab sub-program, ready for use within the main Matlab program.

As with the previous sections, the transfer functions are implemented in Matlab and the system simulated with input frequencies of 2000MHz and 2050MHz. The other ubiquitous constants are as for the Taylor's series case, where the current modulation index was set at 0.4. The extra constants used in the Volterra analysis were set thus [52]; the volume of the active region, $V = 1.56\text{x}10^{-17}\text{m}^3$, the optical confinement factor, $\Gamma = 0.065$, the photon lifetime, $\tau_p = 3\text{ps}$, the optical gain factor, $g = 5\text{x}10^{-12}$, and the gain compression constant, $\varepsilon = 3.5\text{x}10^{-23}$. With these constants declared, and, using the expression for the optical modulation index below:

$$m_{opt} = m_{cur} \frac{|H_1(\omega)|}{|H_1(0)|} \tag{3.29}$$

When calculated, it can be seen that m_{opt} has a value of 0.391. This is close to the m_{cur} value of 0.4, and is much as predicted in section 3.2. Although the m_{opt} term is not explicitly used in the Volterra analysis program, (due to the simplicity of using m_{cur} instead), it is still interesting to note its value.

With the simulations run as in sections 3.1 and 3.2, the resulting outputs for the pre-correction and post correction cases are as seen in Figure 3.5 below.

From these results, first published in May 2002 [74], it can be seen that the 3rd order intermodulation products are reduced by correction from a level of around 60dB below the fundamental to a level of around 110dB below the fundamental. This is a reduction of 50dB, and is consistent with the results of the previous sections, with the initial level of IMD_3 much as was seen for the perturbation case. This would seem to emphasise the preposition that, at modulation speeds in excess of approximately 100MHz, more sophisticated models are required.

(a) Precorrective case.

(b) Final corrected output.

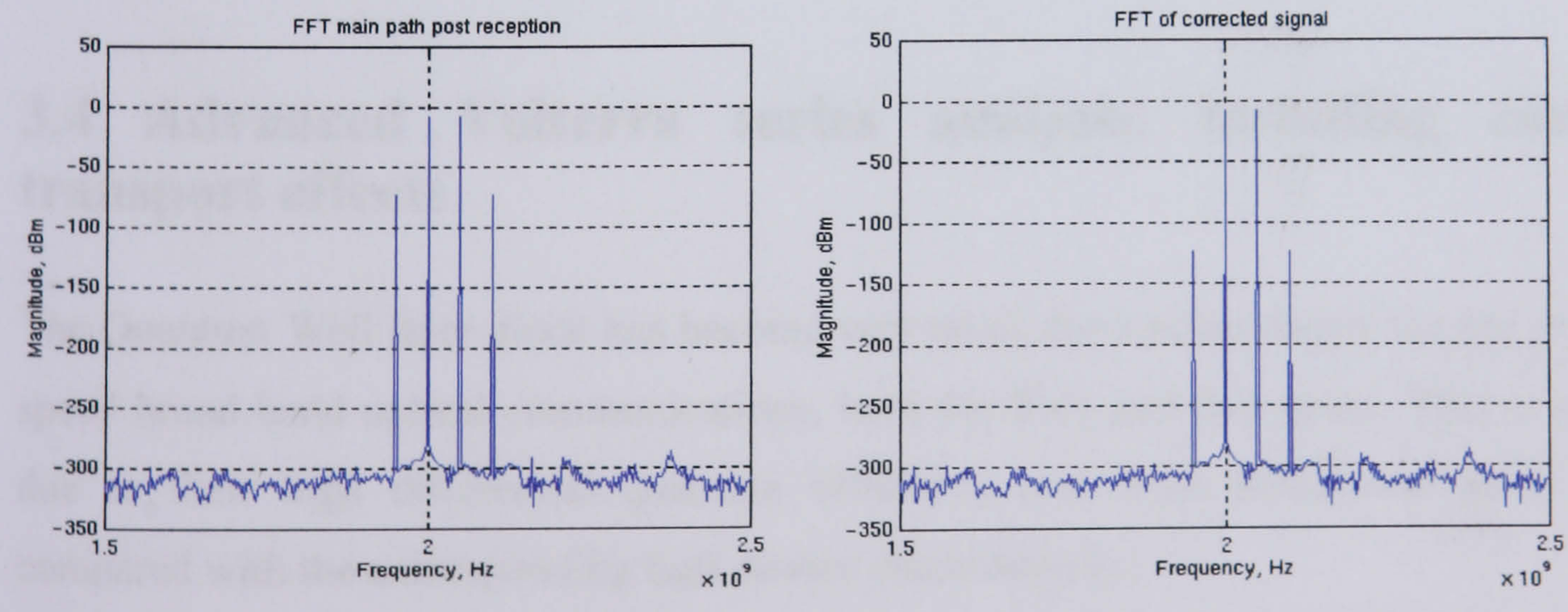


Fig. 3.5 *Results for Volterra series analysis of the optical feedforward non-linearity correction simulation*

As in sections 3.1 and 3.2, the simulations are repeated for a modulation index of 0.2 and 0.6, with the results shown in Appendix C (figures C5 and C6). As with the previous results, the level of IMD_3 is decreased for the $m_{cur} = 0.2$ and increased for $m_{cur} = 0.6$, as expected. The level of IMD_3 compensation for the lower modulation index is also consistent with the previous findings, at around 60dB. However, the level of compensation for the higher modulation index case remains at around 50dB, which is much as in the Taylor's series case, rather than the perturbation analysis case, suggesting that the Volterra analysis holds together better when possible over modulation is used. It is however worth noting that in the graph for the corrected signal for which $m_{cur} = 0.6$, the higher order spikes which are so prevalent in the perturbation case, are just starting to appear (albeit still at a level well below that of the noise floor), indicating that perhaps the modulation depth is reaching the small signal distortion limit.

Some previous studies using Volterra analysis for laser diode modelling have also varied some of the other constants present in the LD expressions. These include the photon lifetime, τ_p, and the gain compression factor, ε. However, the varying of these constants only appears to have a significant influence on the modulation response of the LD, and has little effect on the IMD_3 of the LD [52]. This is proved to be the case, with no noticeable change to the results shown in figure 3.5 upon the varying of τ_p and ε, and hence this aspect is not pursued further in this study.

3.4 Advanced Volterra series analysis, including carrier transport effects

The Quantum Well laser diode has become very much the leading choice for use in high-speed broad band optical communications, both for fibre and free space. This is mostly due to their high differential quantum efficiency and large differential gain, when compared with the corresponding bulk device characteristics.

Analysis of the QW LD characteristics are usually undertaken using the two rate equation model, as in the previous section. However this model is based upon bulk lasers, and is only accurate if certain carrier transport effects are ignored. Investigations have shown

that carrier transport can play an important role in determining the high speed performance, both theoretically and experimentally [75]. The times of carrier capture, τ_{cap}, into the active region (i.e. the quantum well) of the laser diode and carrier escape, τ_{esc}, from the active region, are key parameters that can significantly limit the maximum modulation bandwidth [76]. Also, as the level of distortion produced by the correction path laser is a limiting factor in the achievable compensation, if carrier transport effects have any affect upon this distortion, they must be taken into account.

The inclusion of the two new parameters necessitates the expansion of the two rate equation model (as used in the previous section) to a three rate equation model, with separate carrier density expressions for the different regions within the QW LD. The two regions in question are the separate confinement heterostructure (SCH) and barrier region, and the QW itself. This is illustrated in Figure 3.6 below.

Making the assumption that the theory for a single well, as shown below, holds for the multiple QW device, and, neglecting the contribution from spontaneous emission, the rate equation model for carrier density inside the quantum well $N_{qw,}$ and for carrier density in the SCH and barrier region, N_{sch}, is as follows [76];

$$\frac{dQ}{dt} = Q\left(\Gamma G - \frac{1}{\tau_p}\right) \tag{3.30}$$

$$\frac{dN_{qw}}{dt} = -\frac{N_{qw}}{\tau_n} - \frac{N_{qw}}{\tau_{esc}} + \frac{N_{sch}}{\tau_{cap}} \cdot \frac{V_{sch}}{V_{qw}} - QG \tag{3.31}$$

$$\frac{dN_{sch}}{dt} = \frac{I}{eV_{sch}} - \frac{N_{sch}}{\tau_{cap}} + \frac{N_{qw}}{\tau_{esp}} \cdot \frac{V_{qw}}{V_{sch}} - \frac{N_{sch}}{\tau_n} \tag{3.32}$$

where G is the material gain and can be written as $G = g(1 - \varepsilon Q)(N_{qw} - N_{qw0})$ and N_{qw0} is the carrier density for transparency, V_{qw} is the volume of the quantum wells, V_{sch} is the volume of the SCH and barrier regions, τ_{cap} is the carrier capture time into the quantum well, and τ_{esc} is the carrier escape time from the quantum well. Other symbols are defined

as in section 3.3. Further explanation of the mechanisms behind the parameters τ_{cap} and τ_{esc} in QW LDs, as well as a further four rate equation model, can be found in [52].

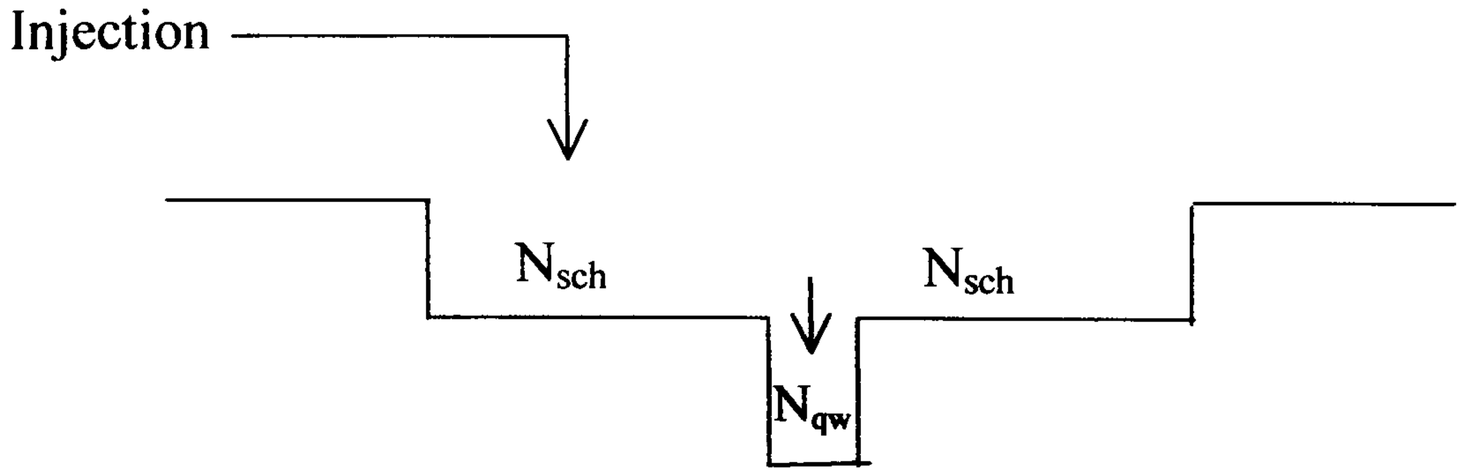


Fig. 3.6 *Schematic of the three rate equation model of carrier transport in QW laser, with the SCH and barrier region in turquoise and the QW region in purple*

As with the previous section, the following mathematical expansion of the rate equations is just a summary, with further explanation to be found in Appendix B2.

Neglecting the recombination of the carrier in the SCH/barrier region, and setting $1/\tau_n$ to zero, as in section 3.3, equations (3.30) to (3.32) can be reduced to;

$$I - I_{th} = \frac{V'}{\Gamma}\left(\frac{dQ}{dt} + \frac{Q}{\tau_p}\right) + \frac{V'(1+R_{rd})}{\Gamma g}\cdot\frac{d}{dt}\left[\frac{\frac{dQ}{dt} + \frac{Q}{\tau_p}}{(1-\varepsilon Q)Q}\right]$$
$$+ \frac{\tau_{cap}V'}{\Gamma}\cdot\frac{d}{dt}\left(\frac{dQ}{dt} + \frac{Q}{\tau_p}\right) + \frac{\tau_{cap}V'}{\Gamma g}\frac{d^2}{dt^2}\left[\frac{\frac{dQ}{dt} + \frac{Q}{\tau_p}}{(1-\varepsilon Q)Q}\right] \qquad \textbf{(3.33)}$$

where, similarly to section 3.3, $I_{th} = qV_{qw}N_{qw0}/\tau_n$ and $V' = qV_{qw}$. Also defined in (3.33) is the capture/escape time ratio, $R_{rd} = \tau_{cap}/\tau_{esc}$, which is also known as the transport factor. It is worth noting that if τ_{cap} is set to zero, resulting in R also being set to zero, equation (3.33) simply reduces to that of equation (3.18).

Following the procedure shown in section 3.3, with $I = I_0 + i(t)$, and $Q = Q_0 + q(t)$, equation (3.33) can be re-written in respect to the driving current as;

$$\begin{aligned} i(t) = A + B_1q(t) + B_2q'(t) + B_3q''(t) + B_4q'''(t) + C_1q(t)q'(t) + C_2q(t)q''(t) + \\ C_2q(t)q''(t) + C_3q(t)q'''(t) + 3C_3q'(t)q''(t) + D_1q'(t)q^2(t) + D_2q'^2(t)q(t) + \\ 0.5D_2q''(t)q^2(t) + D_3q'''(t)q^2(t) + 6D_3q'''(t)q''(t)q(t) + 2D_3q'^3(t) \end{aligned} \quad \textbf{(3.34)}$$

where A, B_1 to B_4, C_1 to C_3 and D_1 to D_3 are defined by the laser intrinsic parameters and are to be found in Appendix B2. Again, following a similar procedure to that shown in section 3.3, the first three-order inverse Volterra transfer functions can be derived, as shown below.

$$G_1(\omega) = B_1 - B_3\omega^2 + j(B_2\omega - B_4\omega^3) \quad \textbf{(3.35)}$$

$$G_2(\omega_1,\omega_2) = -C_2(\omega_1 + \omega_2)^2 + j(\omega_1 + \omega_2)[C_1 - C_3(\omega_1 + \omega_2)^2] \quad \textbf{(3.36)}$$

$$\begin{aligned} G_3(\omega_1,\omega_2,\omega_3) = & -D_2(\omega_1 + \omega_2 + \omega_3)^2 + j[2\,D_1(\omega_1 + \omega_2 + \omega_3) - \\ & 2D_3(\omega_1 + \omega_2 + \omega_3)^3] \end{aligned} \quad \textbf{(3.37)}$$

The required forward Volterra (current to photon) transfer functions, H_1 to H_3, can be derived from equations (3.35) – (3.37), following the same method as described in section 3.3/Appendix B1 with equations (3.24), (B1.28) and (B1.29). The expressions for the fundamental, IMD_2 and IMD_3 are the same as those in equations (3.26) to (3.28).

At this point it is worth noting that, although the above model is improved upon that used in section 3.3, there is still scope for further accuracy. If the SCH is long, which is common among MQW lasers, then the carrier transport analysis should include carrier diffusion within the SCH region, and local carrier capture and escape within the quantum wells.

A full analysis of these additional parameters requires a four rate equation model, an example of which can be found in Liang, section 4.3.3 [52].

However, under direct current modulation, the small signal response for the LD has been found to be similar to that of the three rate equation model if certain transformations to

the parameters of R_{rd} and T_{cap} are performed. These transformations can be seen in various literature, including Liang, [52], [77], and hence will not be shown here. However it will be assumed that the values subsequently used for R_{rd} and T_{cap} will include the transformation scale-ups.

These new expressions have been, as in section 3.3, implemented in a Matlab sub-program and were then included within the main Matlab simulation program. The values for R_{rd} and τ_{cap} were shown to vary from 1 to 5 for R_{rd}, and 1ps to 60ps for τ_{cap} [77], [78]. Previous studies have shown that the value of R_{rd} has the more significant affect on the level of IMD than the value of τ_{cap}, with an increasing value of R_{rd} resulting in an increasing level of IMD. However, the value of τ_{cap} affects IMD in the opposite manner, with the larger values resulting in a reduced IMD.

The results of the simulations, including those shown in Figure 3.7 below, confirm these previous findings, with the highest increase in IMD_3 coming with R_{rd} at a maximum of 5, combined with a small (effectively minimum) value of τ_{cap} of 1ps. The diagrams in Figures 3.7, C7 and C8 show the simulations for values of R_{rd} and τ_{cap} of 3 and 30ps, 1 and 60ps, and 5 and 1ps respectively (i.e. scenarios showing lowest, highest and intermediate levels of IMD). In all cases the modulation index was set at 0.4. The first point to note is that the first pair of diagrams, for which $R_{rd} = 1$, are almost identical to the standard Volterra case, shown in Figure 3.6, (see figure C7). In the figure 3.7 pair of diagrams, when $R_{rd} = 3$, the level of IMD_3 has risen to around 50dB below carrier and it can be seen. that the level of cancellation has started to reduce, to that of around 40dB. This trend is further continued in the case of the third pair of diagrams, where the IMD_3 is the highest, at only around 45dB below carrier as shown in figure C8. Here there does appear to be a reduction in the amount of distortion cancellation to around 30dB, which is however still significant. It was suggested in Chapter 2, section 2.5.3, that the reduction in IMD_3 cancellation is limited mainly by the IMD_3 introduced by the correction path laser.

(a) Precorrective case.

(b) Final corrected output.

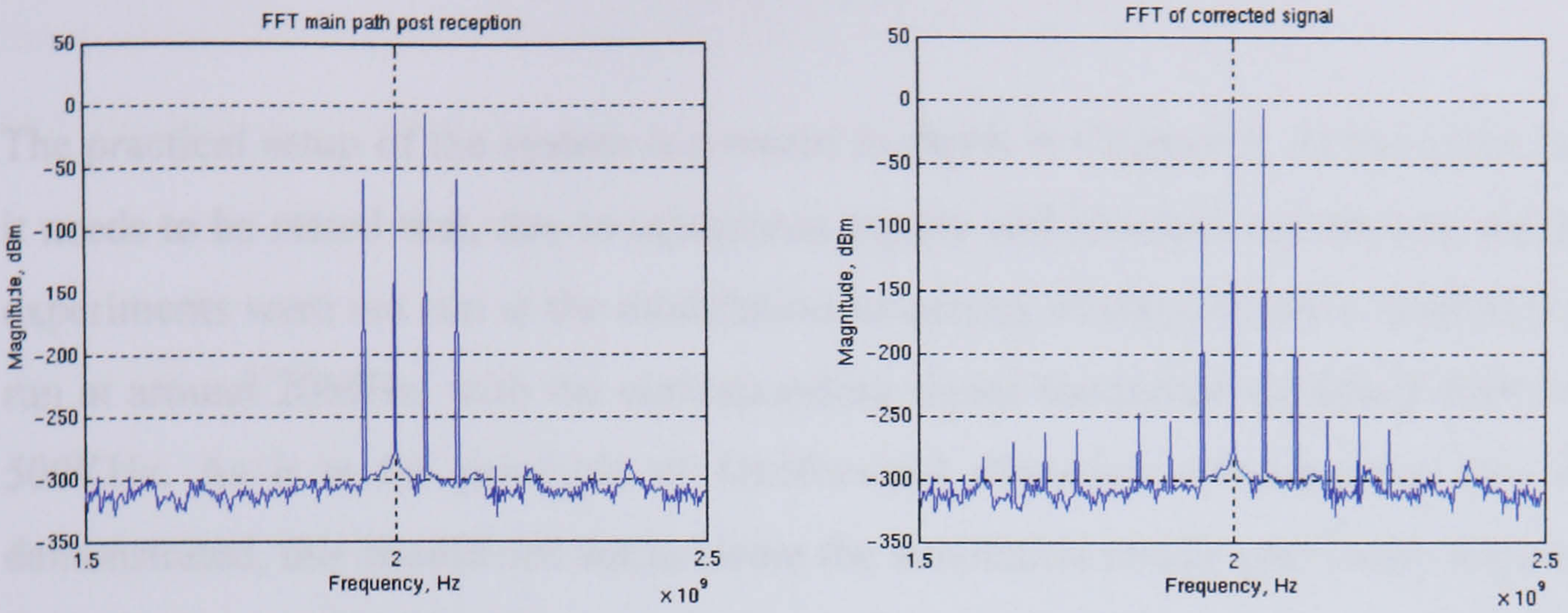


Fig. 3.7 *Results for Volterra series analysis of the optical feedforward non-linearity correction simulation, including carrier transport effects*

The results shown here would appear to confirm this, as when the overall level of $IMD_3$is high, the cancellation is reduced. However, while this is not unexpected, there does appear to be a point when the level of IMD_3 is within 50dB of the carrier, where the level of distortion compensation reduces more significantly. This scenario was also apparent in the over modulated perturbation simulation, shown in figure C4. The simulation for R_{rd} = 5, τ_{cap} = 1ps is repeated with a modulation index of 0.6 with the results shown in figure C9. Although the IMD_3 is now only 40dB below carrier, the IMD_3 compensation holds at around 30dB, whilst the higher order spikes (which first appear in the 2nd diagram pair of Figure 3.7) are now at a level where they are close to breaking through the noise floor, albeit still well below a level that would cause any noticeable interference to the desired signal. This would appear to indicate that, although the main IMD compensation limitation is the level of IMD in the correction path signal, there will be other sources of limitation, and indeed once the distortion becomes too extreme, the analysis used above might become unreliable.

Overall the results of the simulations in this section show compensation levels far more akin to those expected in a practical experiment. This would appear to indicate that in the light of the limiting IMD_3 introduced by the correction path laser, the advanced Volterra analysis does indeed yield more accurate simulations.

3.5 Simulations at low frequency

The practical setup of the system is covered in depth in Chapter 5. At this point however, it needs to be stated that, due to equipment supply and financial constraints, the practical experiments were not run at the modulation frequency of around 2GHz. Instead they were run at around 20MHz, with the corresponding signal bandwidth dropping from 5MHz to 500KHz. As it is the principle of feedforward distortion compensation that is to be demonstrated, this should not detract from the simulation results previously acquired.

Nevertheless, replicating the system simulations at these lower frequencies will give a more direct comparison with the practical system, as well as providing further insight into how the differing LD models behave at different frequencies.

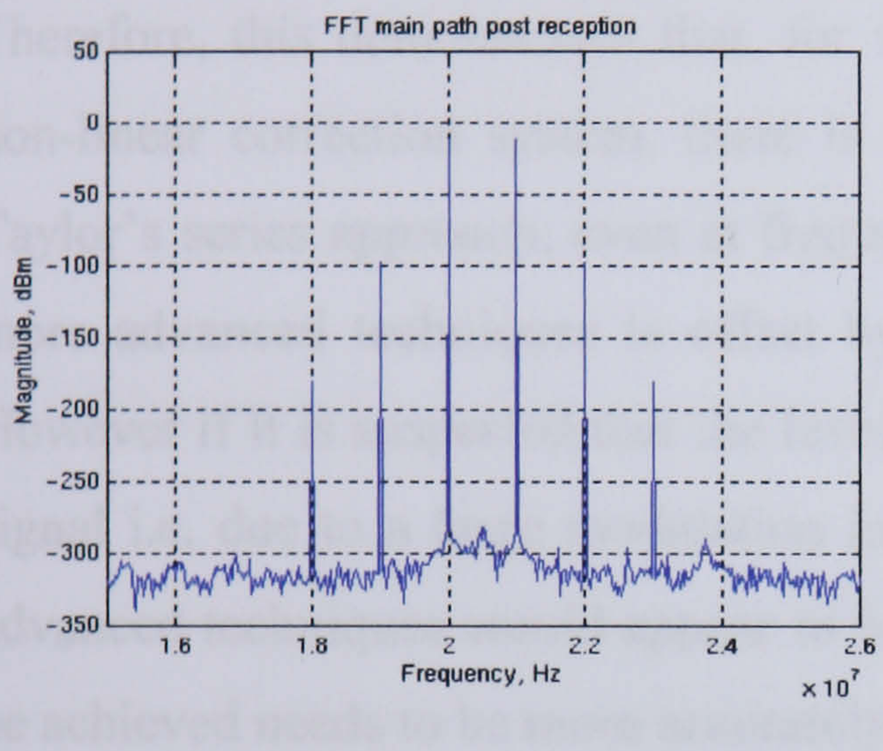
In this section, the two models that are used are the Taylor's series and the advanced Volterra analysis. As previously stated, a laser diode starts to exhibit significant memory properties in its analysis once the modulation frequency reaches approximately 100MHz. Below this frequency, though the memory effects may still be there, they do not exert such influence on the laser diode response. This means that the Taylor's series analysis should give a reasonably accurate idea of the amount of distortion compensation possible. The advanced Volterra analysis is included for comparison, to see how it changes from the high frequency model, where it gives the most accurate simulations. Keeping with an optical modulation index of 0.4, the two models were adapted for two tone input frequencies of 20MHz and 21MHz. The rest of the system's DC characteristics remained as for the high frequency case.

From the top pair of diagrams in Figure 3.8 below, it can be seen that the levels of IMD_3 for the Taylor's series model, both pre and post correction, are identical to those shown in Figure 3.1. This to be expected as the Taylor's series analysis has no frequency dependent amplitude components.

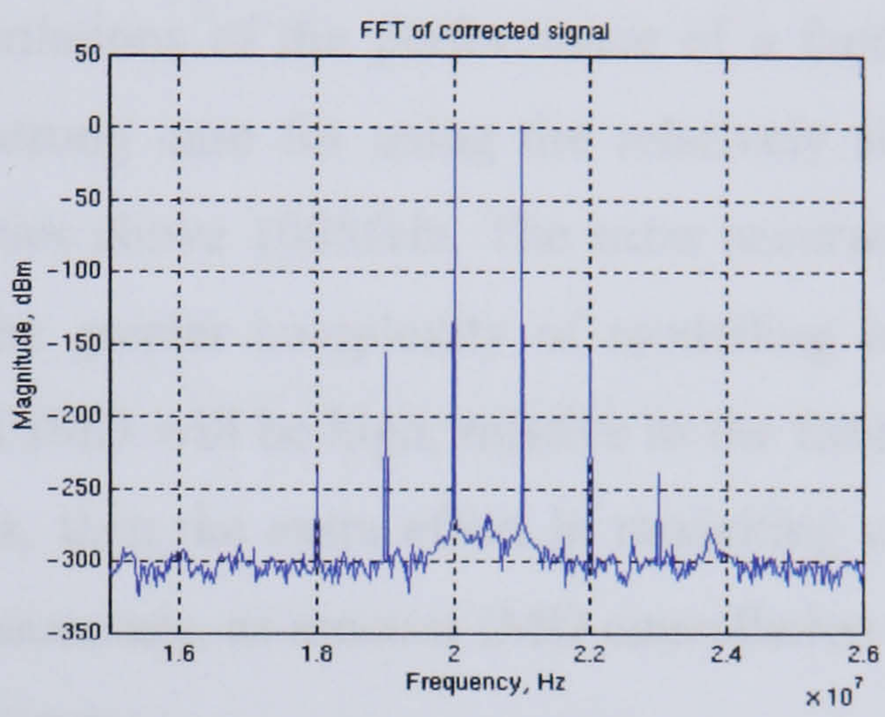
This is not the case with the Volterra model. Using the intermediate level settings described in section 3.4, it can be seen from the bottom pair of diagrams in Figure 3.8 that the pre-correction level of IMD_3 is much reduced from the case in Figure 3.7, to around 110dBc, and the level of compensation achieved is an optimistic 190dBs.

These results would appear to confirm the theory that at lower frequencies, the use of the more simplistic Taylor's series analysis is sufficient for the model of a feedforward distortion correction system.

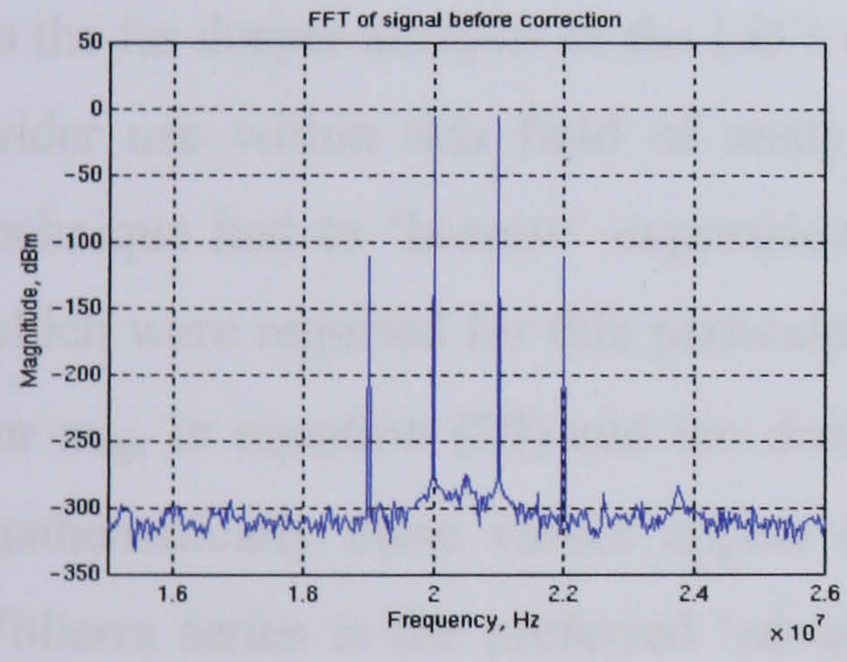
(a) Taylor's series precorrective case.



(b) Taylor's series final corrected output.



(c) Volterra series precorrective case.



(d) Volterra series final corrected output.

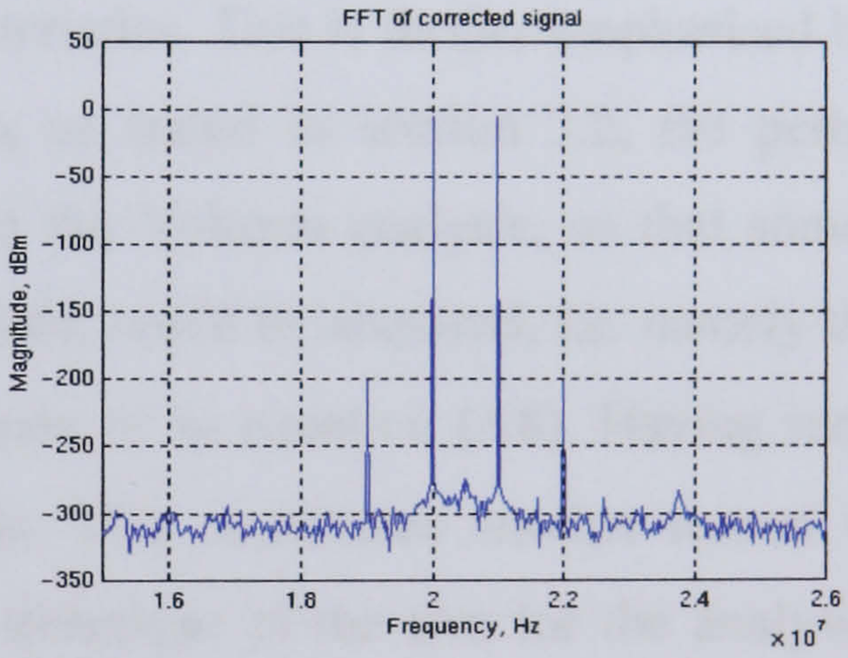


Fig. 3.8 *Results of the Taylor's series and Volterra series simulation at modulation frequencies of 20MHz and 21MHz*

3.6 Summary and conclusions

In this chapter the free space feedforward linearisation system introduction in Chapter 2 was modelled in Matlab, with special attention shown to the laser diode characteristics.

These characteristics were modelled using the relatively simple Taylor's series, and the more mathematically involved techniques of perturbation and Volterra series analysis. Although the more advanced techniques do appear to give a more accurate model of LD behaviour, the interesting, and important point to note is, that as long as the level of non-linearity, or IMD_3 in this case, is not too high in comparison to the fundamental signal, the level of IMD compensation predicted is much the same.

Therefore, this demonstrates that, for simulations of the performance of a feedforward non-linear correction system, there is a strong case for using the relatively simplistic Taylor's series approach, even at frequencies above 100MHz. The extra accuracy of the more advanced techniques is offset by the greater complexity of modelling involved. However if it is suspected that the level of IMD will be high, relative to the fundamental signal i.e. due to a large modulation index, then the extra effort in modelling using the advanced techniques would appear to be necessary, as amount IMD cancellation that can be achieved needs to be more accurately known.

Of these two techniques, the Volterra series has been shown to be the more accurate, due to the far deeper analysis of the LD's characteristics. This is further emphasised by its far wider use within this field of study. Also, as stated in section 3.2, the perturbation technique had to 'borrow' expressions from the Volterra analysis, so that some values which were required for this particular analysis, could be acquired, i.e. namely the value for m_{opt} in equation (29) and the damping rate, γ, in equation (A8). Having stated this, mathematically these values appear realistic. This constitutes another reason why the Volterra series is the preferred 'advanced' technique of the two for the analysis of the free space, feedforward non-linear correction system.

Henceforth, any further simulations undertaken within this study for the said system will include only the Taylor's series and Volterra series.

Chapter 4 – Noise Analysis

In the world of communications, no system can be integrated into a practical environment without some level of noise analysis. Therefore this chapter develops an analysis of the feedforward non-linear correction system with regards to its noise performance. The noise models are written and simulated in Matlab.

4.1 Standard noise analysis (with 1:1 coupling ratio)

As previously stated, the feedforward technique has been shown to offer some reduction in noise, as well as correcting for non-linearities. In fact, it is only the noise sources present in the main modulation path, as shown in figure 4.1 below, that can be corrected for, as it is the only noise in the system that is inverted and re-combined in anti-phase with itself. If it is assumed that any noise coming from the signal source is negligible, then, the noise that is corrected is solely the main path laser diode's relative intensity noise, or RIN, (as defined in section 2.2). This leads to problems in using the conventional noise analysis technique, described in Chapter 2, of referring to all the system's noise sources as equivalent input noise, or EIN. If all the noise sources are referred back to the system inputs as EINs and then propagated back through the feedforward system, the recombination will result in a cancellation of additional noise sources to that of the main path laser RIN, namely, any noise other sources in the correction path.

The solution to this is to employ a more traditional 'end-to-end' mathematical model of the noise, by the addition of the noise sources (as defined for each component in Chapter 2) as they propagate through the system [79], while ignoring the modulation signal. However since each noise source is considered white, and hence uncorrelated, they cannot be simply added in such a manner. Instead, the sum of root squares method has to be employed. In practical terms, this means the addition of the noise spectral densities for

each source. For this work the noise current spectral densities are chosen. In addition, all the sources are considered equally random throughout the frequency range and hence are calculated for a unit bandwidth. This results in being able to treat the final noise outputs as rms signals, which considerably simplifies the analysis within Matlab.

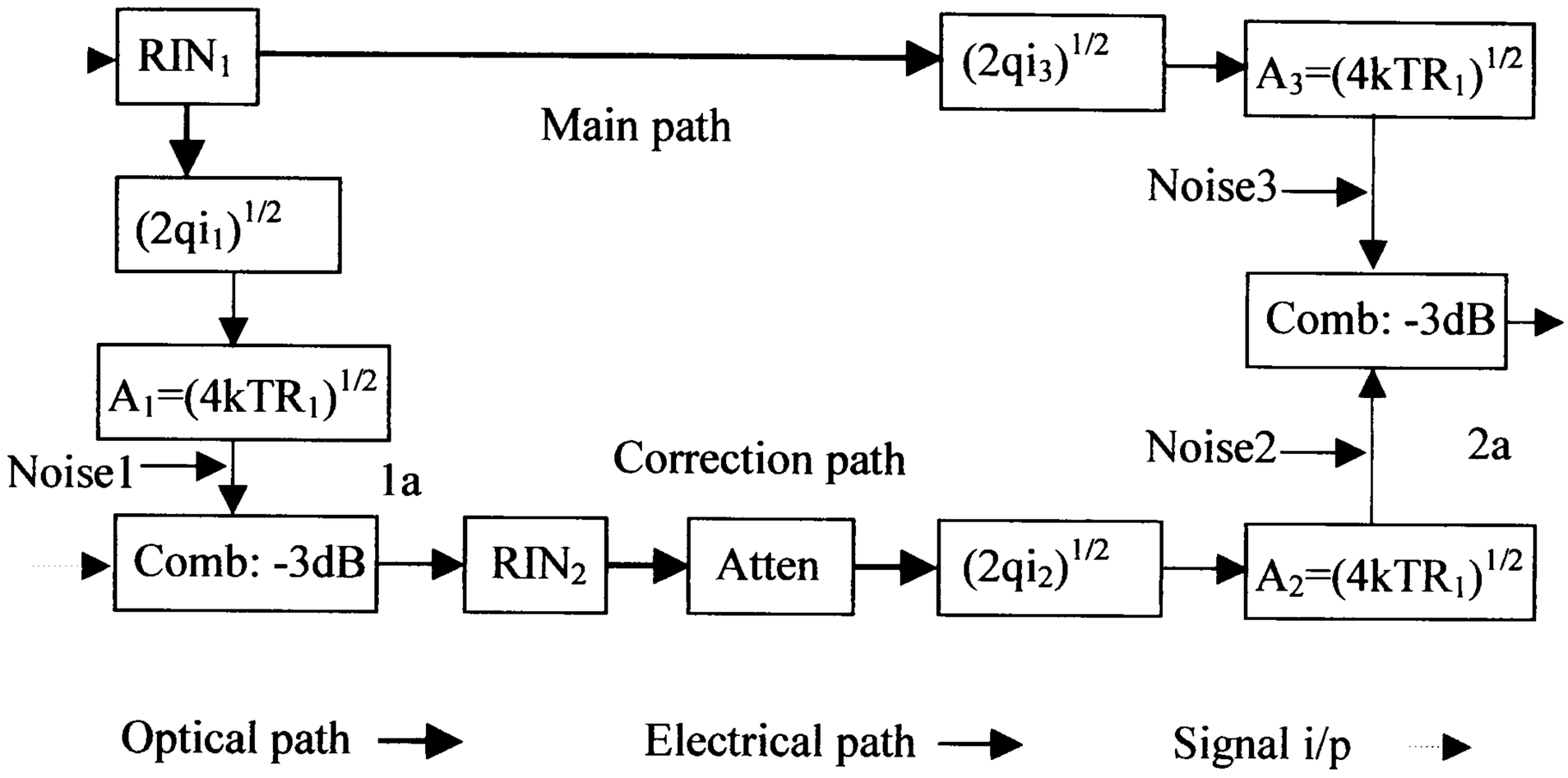


Fig. 4.1 *Unit bandwidth noise source schematic of the feedforward correction system of figure 2.10*

From figure 4.1 it can be seen that there are the following three different type of noise source; shot noise, thermal noise and RIN, associated with the photo diodes, amplifiers and laser diodes respectively. In addition to the noise sources, Figure 4.1 also includes the two combiners from the original system, each with a 3dB insertion loss, and an attenuation block in the correction path. This attenuation block represents the losses that are incurred by the noise modulating the correction path laser diode, LD_2. These factors do not affect the output of the main laser diode, LD_1, as the noise produced is in the form of RIN and, as defined as a ratio of two square powers, is unaffected by the attenuated over the free space link section.

Also shown in the above diagram are the Noise points 1, 2 and 3, and are included to help clarify the noise analysis as it builds throughout the system.

The noise sources of the LDs, PDs and amplifiers in figure 4.1 are all of differing quantity, being a ratio, a current and a voltage respectively. For the addition of these sources to occur they need to be converted to equivalent current spectral densities, Si_{RIN}, Si_{shot} and Si_{amp}. This is shown in equation (4.1) below, where the noise at point 1 from figure 4.1 is calculated as;

$$\begin{aligned} \text{Noise1}_{A^2/Hz} &= G_1 (Si_{RIN1} + Si_{shot1} + Si_{amp1}) \\ &= G_1(RIN_1 I_1^2 + 2I_1 q + 4kTF_1/R_1) \end{aligned} \qquad \textbf{(4.1)}$$

Where I_1 is the direct current in PD_1, q is the electron charge, k is the Boltzmann constant, T is the absolute temperature and F_1 is the noise figure and R_1 the (nominally matched) resistance for amplifier$_1$. G_1 is the power gain of amplifier$_1$.

Noise1 now undergoes a reduction of 3dB due to the signal combiner. This noise signal now modulates LD_2. As the modulation signal Noise1 will be reduced by the internal efficiencies of the laser diode and the receiving photodiode, as well as attenuation from the air interface. Referring to table 2.1 in Chapter 2, it can be seen that for a 100m link, only the severest weather conditions will have any great effect. For a moderate haze, the attenuation is around 3dB/km, which translates as 0.3dB for a 100m link. However, as the equivalent electrical loss of any optical loss is twice the level, this becomes 0.6dB electrical loss. With internal efficiencies of the LD and PD of around 0.35, this gives a modulation reduction coefficient, η_{link}, for the LD_2 to PD_2 link of around 0.1. (Note: When fed into the following expressions for current spectral density, η_{link} must be squared).

Following the modulation signal through to point Noise2 in figure 4.1, the expression develops as shown;

$$\text{Noise2}_{A^2/Hz} = G_2 (\eta_{link}^2 \, \text{Noise1} + 2I_2 q + RIN_2 I_2^2 + 4kTF_2/R_2) \qquad \textbf{(4.2)}$$

Where I_2 is the direct current in PD_2, RIN_2 is the relative intensity noise of LD_2, and F_2 is the noise figure, R_2 resistance and G_2 the power gain for amplifier$_2$.
Noise3 is calculated in an identical way to Noise1 as shown below;

$$\text{Noise3}_{A^2/Hz} = G_3(RIN_1 I_3^2 + 2I_3 q + 4kTF_3/R_3) \qquad \textbf{(4.3)}$$

Where I_3 is the direct current in PD_3, and F_3 is the noise figure, R_3 resistance and G_3 the power gain for amplifier$_3$.
Noise3 is also the noise output for the uncorrected system, i.e. if the error correction path is switched off. The corrected noise output, Noise_{final} is then sum of Noise2 and Noise3, divided by two to account for the combiner, as shown;

$$\text{Noise}_{final,\ A^2/Hz} = (\text{Noise3}_{A^2/Hz} + \text{Noise2}_{A^2/Hz})/2 \qquad \textbf{(4.4)}$$

Within Matlab, Noise3 and Noise_{final} are then put across a matched load resistance and adjusted to give the noise power outputs in dBm/Hz, (as shown in the results below).

If equations (4.1) to (4.3) are substituted into equation (4.4), it can be seen that there are two RIN_1 terms. These terms are correlated and will be of opposite sign, due to the error cancellation process, and thus, can be written as,

$$G_1 G_2 \eta_{link}^2 RIN_1 I_1^2 = G_3 RIN_1 I_3^2 \qquad \text{or} \qquad G_1 G_2 \eta_{link}^2 I_1^2 = G_3 I_3^2 \qquad \textbf{(4.5)}$$

when the feedforward correction is correctly adjusted and thus, RIN_1 cancelled. If, in equation (4.5), it is assumed that the values of I_1 and I_3 will of a similar order, it can be seen that maximum RIN cancellation (timing complexities aside) depends largely upon careful signal amplitude adjustment via the amplifier gains. This is same as for the cancellation of the intermodulation distortion discussed in the previous chapter.
If equation (4.5) is then substituted into equation (4.4), the resulting expression would then give the noise limit in the system and when compared with the noise source RIN_1, the reduction in noise available can be seen.

For the Matlab simulations, the values of the parameters introduced in equations (4.1) – (4.5) are as follows:

$R_1 = R_2 = R_3 = 50\Omega$

$RIN_1 = -120dB/Hz$

$RIN_2 = -120dB/Hz$

$G_1 = G_2 = G_3 = 14dB$

$F_1 = F_2 = F_3 = 3.8dB$

$q = 1.6 \times 10^{-19}$ C

$k = 1.381 \times 10^{-23}$ J/K

$T = 290$ K

$\eta_{link} = 0.1$

$I_1 = 1mA$

$I_2 = 0.5mA$

$I_3 = 0.8mA$

Where G_1, G_2, G_3, F_1, F_2 and F_3 are the gains and noise figures for the three photo receivers amplifiers that are used in the practical set up, (see Chapter 5). In the analysis, the input values for I_1, I_2 and RIN_1 are varied and the outputs for the corrected and non-corrected systems are compared. RIN_1 is chosen as a variable due to it being the noise source that is cancelled. I_1 is chosen due to it being a potentially large front-end variable noise source and its prominence in equation (4.5). I_2 is the major variable noise source linked with the corrected path laser diode, i.e. the larger the correction signal, (which is desirable for longer distance transmission) the greater the value of I_2, and hence, its effect is of interest.

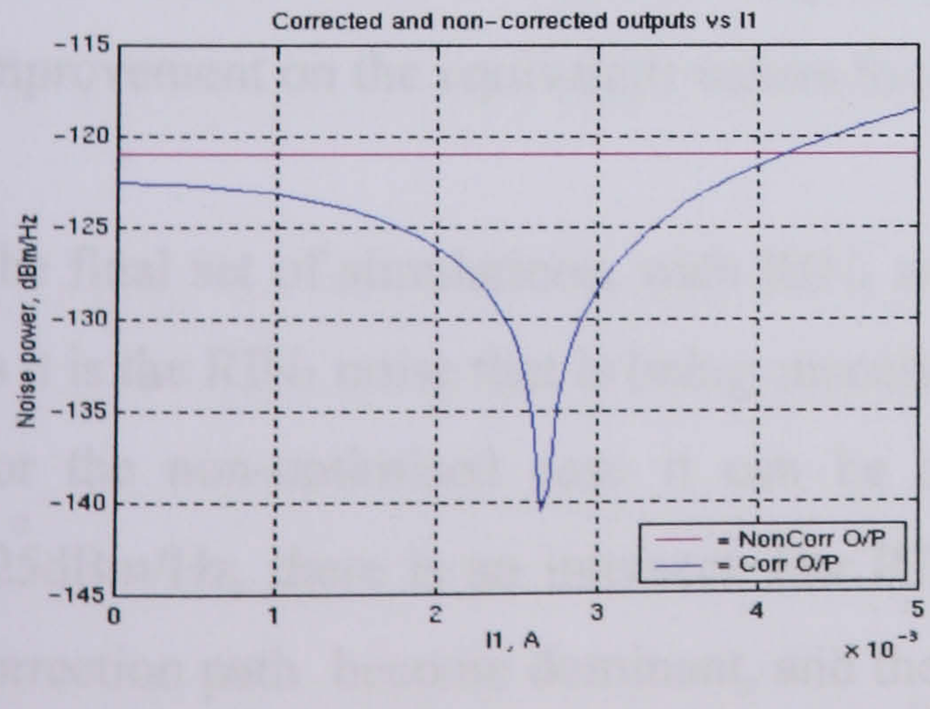
When the above numbers are entered into equation (4.5), the left-hand side turns out to be 2.56 times smaller than the right hand side. Therefore, extra amplification of 2.56 or 4.1dB is needed for the left-hand side, which translates practically to the correction path in figure 4.1, for the maximum cancellation of RIN_1 to occur.

This leads to the question of where in the correction path should the additional amplifier be situated. The two options are for it to be placed post the two existing amplifiers, (denoted by A_1 and A_2 in figure 4.1), becoming amplifier 1a or 2a.

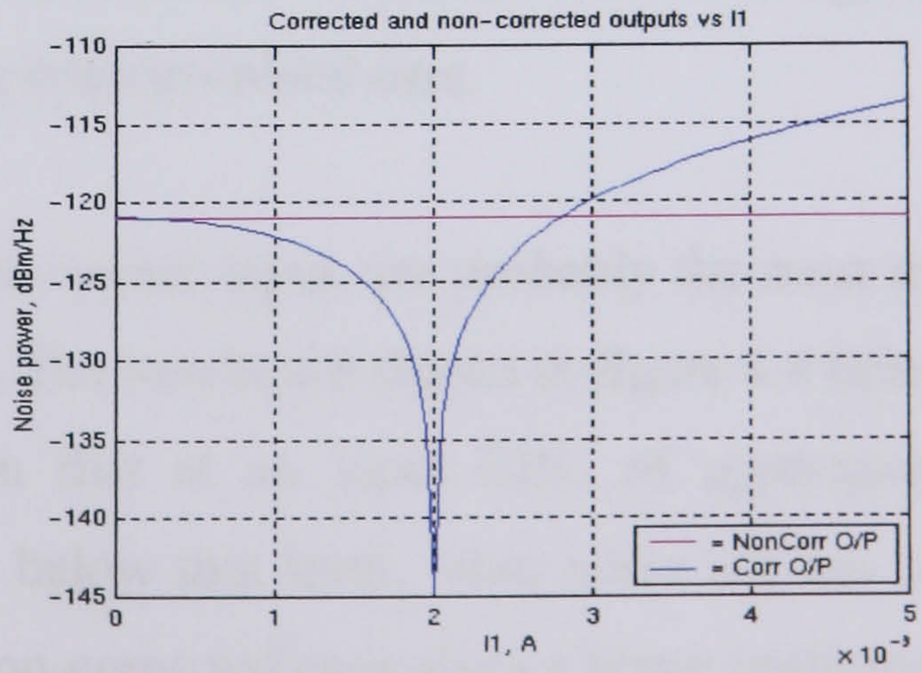
Generally in communications systems the main amplification is placed at the front end of the system. This can be seen from the above equations, where if G_{11} ($G_1 + G_{1a}$) is sufficiently large, then any noise generated in amplifier A_2 would be negligible in comparison, which backs up this assumption. The following results will show the simulations also show this to be the case.

Figure 4.2 below shows the results of varying I_1 against the corrected and non-corrected outputs. As can be seen, the graphs show a sharp dip in the corrected output case. This is due to the fact that all the noise sources are effectively treated as rms signals rather than randomly distributed signals. Therefore, as I_1 is stepped up from 0 to 5mA, there comes a point when equation (4.5) breaks down due to the two terms being of opposite sign. Also, though the optimised cases are shown in the second and third graphs, in reality they are not optimised due to I_1 changing. Therefore the optimised diagrams are only really applicable for when I_1 is set at 1mA, where the difference between the two outputs is 2.2dB, 1.2dB and 3.5dB for the non-optimised, and optimised amplifier position 2a and position 2b respectively. This indicates that the 'front end' positioning of the optimisation amplifier gives the best noise compensation results, though it can be concluded, from figure 4.2, that varying I_1 against the outputs does not give very satisfactory results.

(a) Non-optimised case



(b) Optimised case, position 2a



(c) Optimised case, position 1a

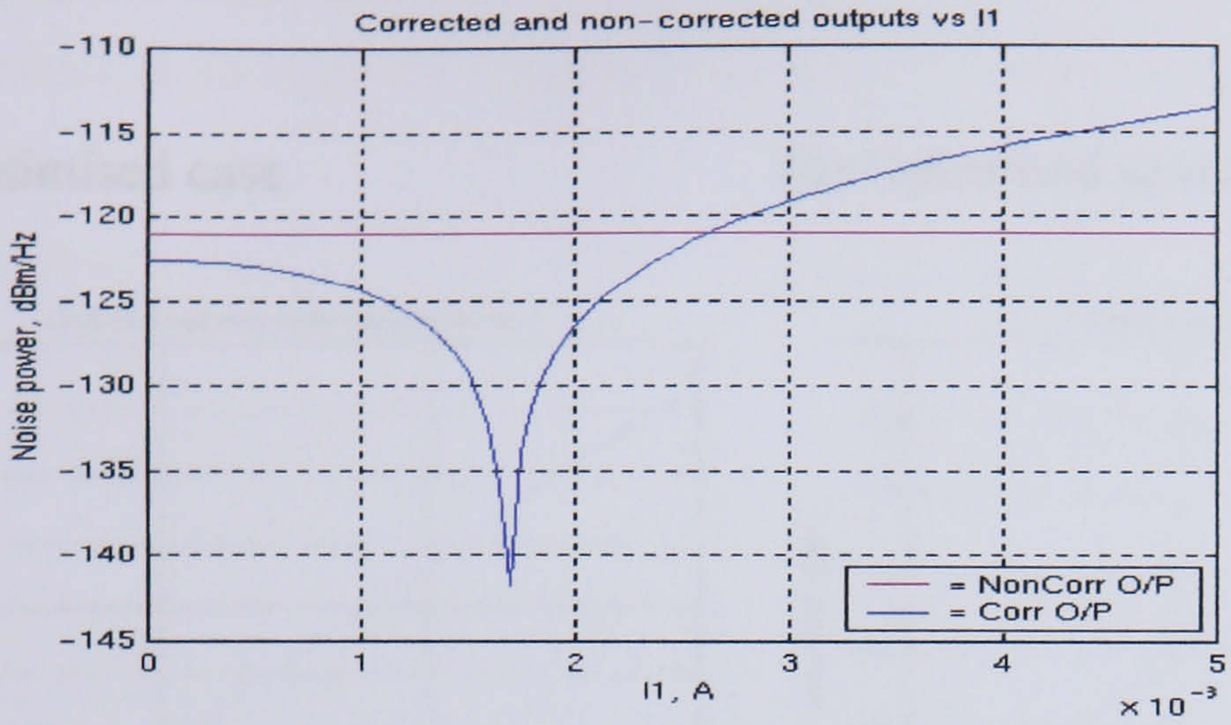


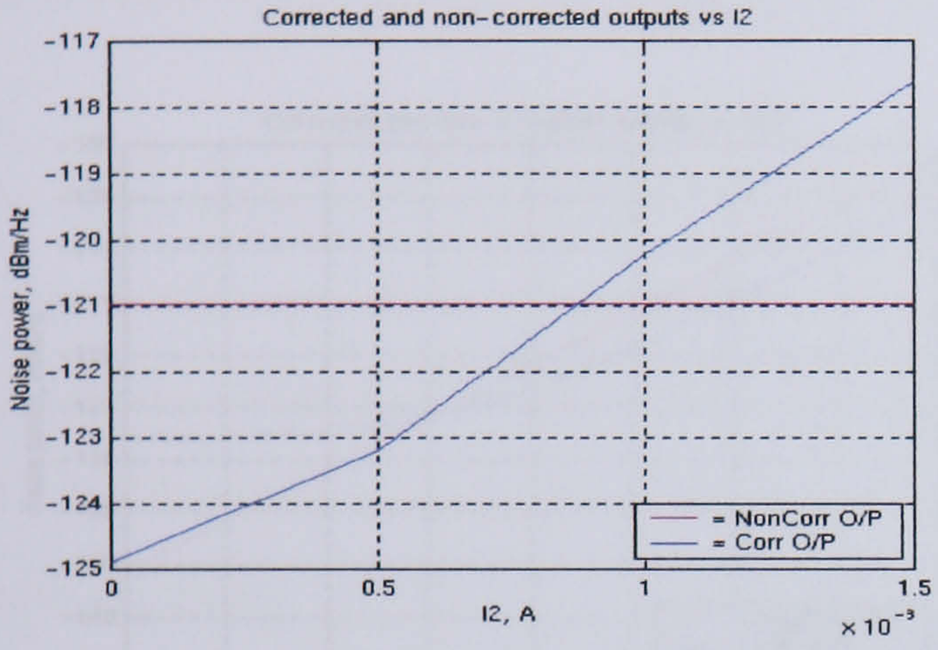
Fig. 4.2 *System noise outputs against variation of I_1*

With I_2 being the varied input, the results are far more satisfactory, as shown in figure 4.3. The two parameters to note in the three graphs are, the value of I_2 at which the corrected output intersects with the non-corrected output, and the difference between the outputs when I_2 is at its designated value of 0.5mA. As can be seen, the non-optimised case actually performs better than the optimised Amp_{2a} case, with an intersect I_2 at approximately 0.87mA and a difference value of around 2.2dB, compared to an intersect of 0.63mA and a difference value of 0.6dBm. It is not too surprising that the optimised case here suffers, as it will have the most significant effect on the signal produced within PD_2, which are dependent upon I_2. The analysis is more encouraging for the preferred position of the optimisation amplifier, Amp_{1a}. Here the intersect is just below 1mA and the difference value of the two outputs for I_2 at 0.5mA is 3.5dB. This is a significant improvement on the equivalent values for the non-optimised case.

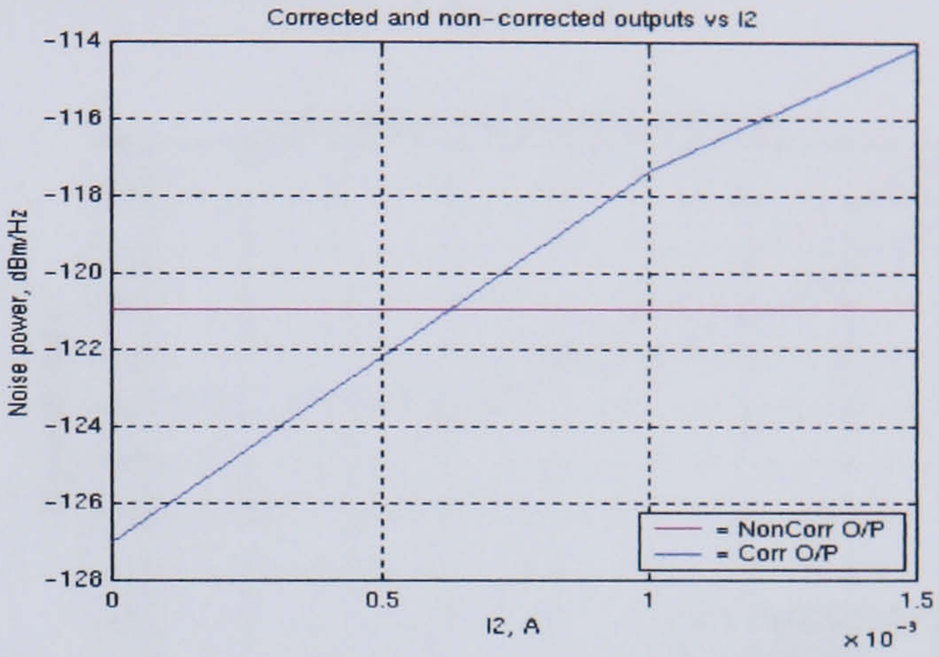
The final set of simulations, with RIN_1 as the varied input, are probably the most useful as it is the RIN_1 noise that is being cancelled. The results are shown in figure 4.4 below. For the non-optimised case it can be seen that at an input RIN_1 of approximately 125dBm/Hz, there is an intersect. For RIN_1 below this level, other noise sources in the correction path become dominant, and the non-corrected case gives a better performance. However above this level, the corrected system performs better with a steady

improvement of around 4dB being achieved, and, importantly, a 2.5dB improvement at an input RIN_1 of 120dBm/Hz.

(a) Non-optimised case



(b) Optimised case, position 2a



(c) Optimised case, position 1a

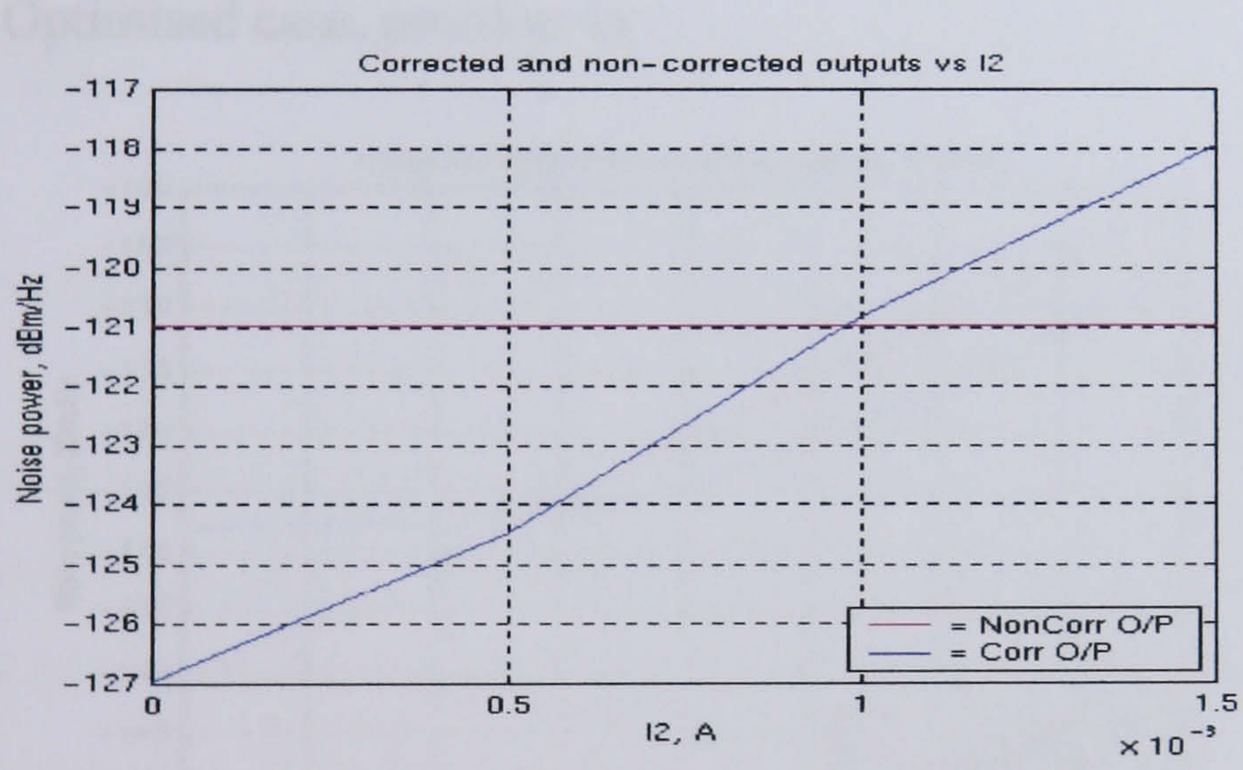


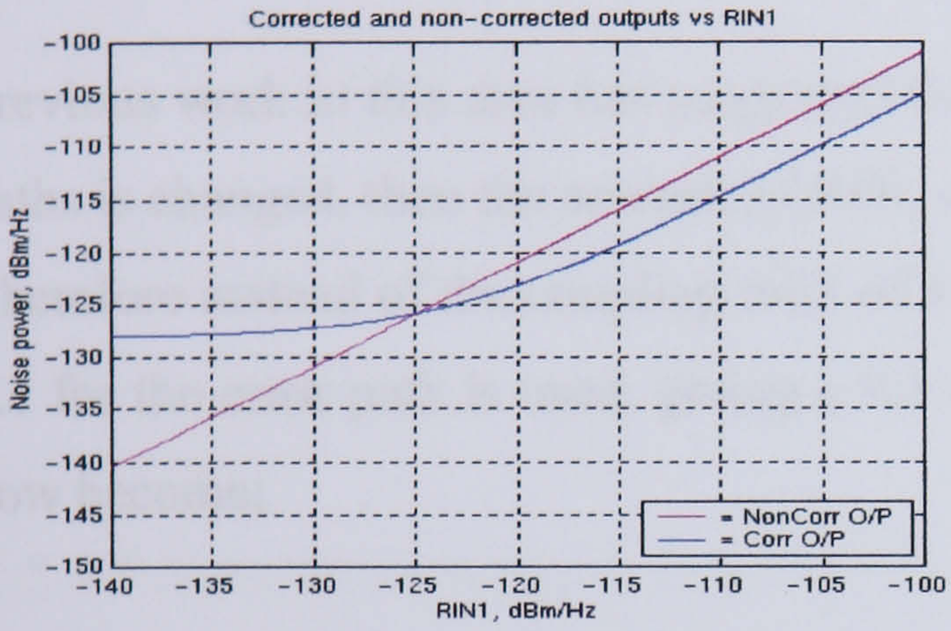
Fig. 4.3 *System noise outputs against variation of I_2*

Following on from the above results, the performance of the correction system is degraded with the optimisation amplifier placed in position 2a, Amp_{2a}. The high input RIN_1 performance has improved to almost 6dB, however both the output intersection and 120dBm/Hz input characteristics are significantly worse.

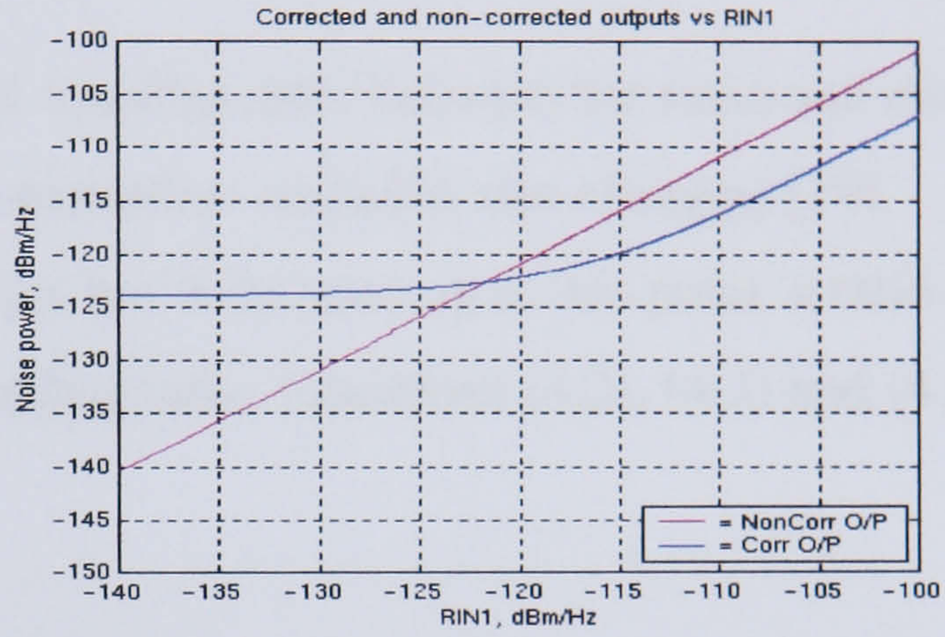
With the optimisation amplifier in position 1a however, there can be seen an improvement in all three characteristics when compared to the non-optimised case. There

is only a small improvement in the intersection, of 1dB. However the high input RIN_1 performance has improved to around 7dB. Most important, is the improvement of the noise difference at an input of 120dBm/Hz which has improved by 4dB.

(a) Non-optimised case



(b) Optimised case, position 2a



(c) Optimised case, position 1a

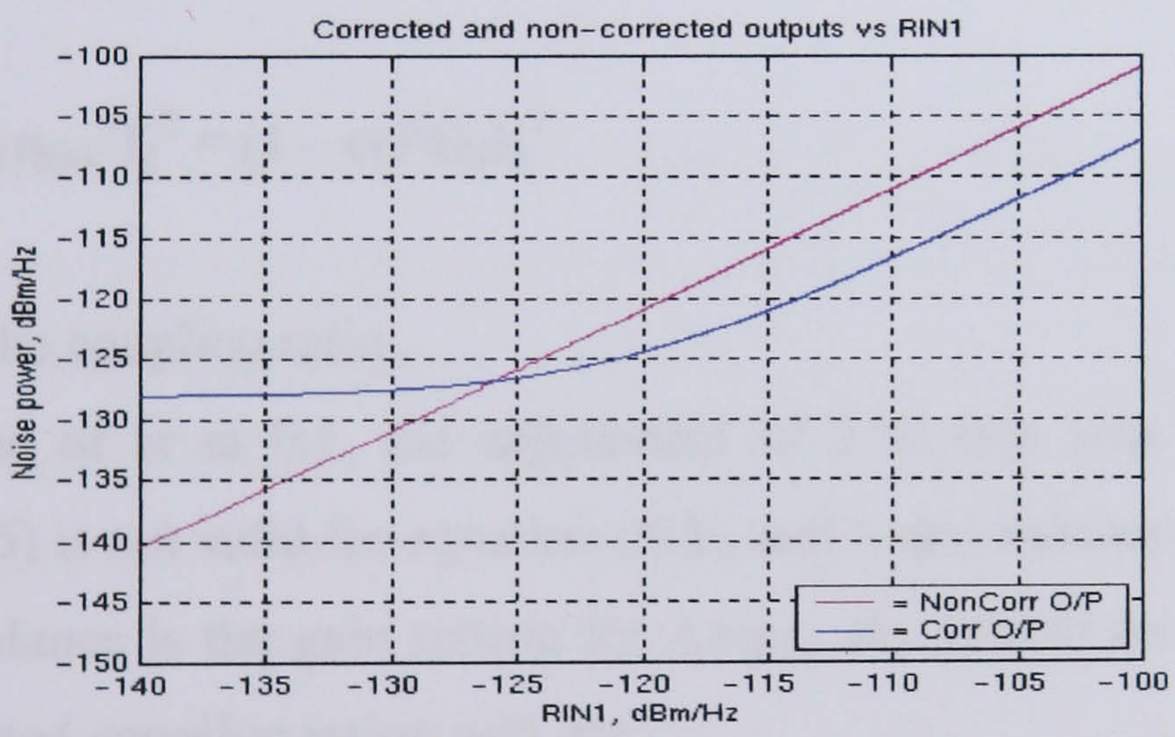


Fig. 4.4 *System noise outputs against variation of RIN_1*

From the above results, it can be reasoned that the best position for the optimisation amplifier is that of 1a, immediately following the photo diode/receiver amplifier circuits. It is important to remember that this amplifier, Amp_{1a} is an inverting amplifier. The above results also suggest that a small drop in noise for the system of around 4dB, can be expected. In the practical case, further amplification and attenuation may be needed

throughout the system, which will have a further impact upon the noise performance. This is covered in Chapters 5 and 6.

4.2 Advanced noise analysis (with a 9:1 coupling ratio)

Previous work in this area has suggested that if coupling ratio between the main and error paths is changed, then the amount of RIN_1 compensation available also changes, [79]. Therefore instead of the coupling ratio of 1:1 that has been used up to this point, a ratio of 0.1 for the error path is used, giving a 9:1 coupling ratio. Equations (4.2), (4.3) and (4.5) now become;

$$\text{Noise2}_{A^2/Hz} = G_2 (\eta_{link}{}^2 \text{Noise1} + 2I_2q + cr^2RIN_2I_2 + 4kTF_2/R_2) \quad \textbf{(4.6)}$$

$$\text{Noise3}_{A^2/Hz} = G_3((1 - cr)^2RIN_1I_3{}^2 + 2I_3q + 4kTF_3/R_3) \quad \textbf{(4.7)}$$

$$G_1G_2\eta_{link}{}^2I_1{}^2 = (1 - cr)^2G_3I_3{}^2 \quad \textbf{(4.8)}$$

where cr is the coupling ratio.

With a value of cr at 0.1, the adjustment of 2.56 that was calculated for balance in equation (4.5) is not valid for equation (4.8), and a new balance of 2.06 is calculated.

This new balance is the gain setting for Amp_{1a}, the system for which all the simulations for the adjusted coupling ration will use.

Running the simulations for the variable inputs of I_2 and RIN_1 only, the results can be seen in figure 4.5 below.

By comparing these results with the equivalent results in figures 4.3 and 4.4, it can be seen that there is a significant improvement in the noise characteristics.

Taking the I_2 input case first, the intersect between the two output is off the scale and the difference at an input value of 0.5mA is up from 3.5dB to nearly 6dB.

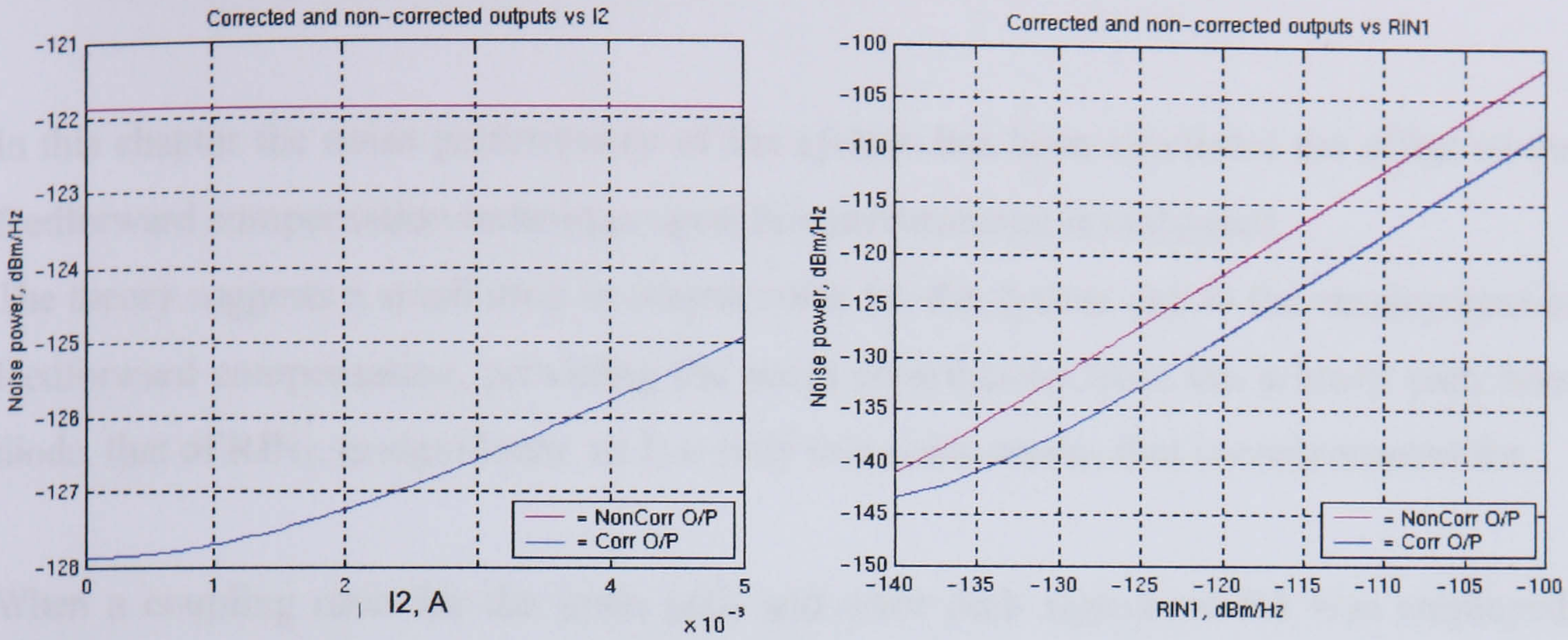


Fig. 4.5 *System noise outputs against inputs of I_2 and RIN_1 for a coupling ratio of 0.1*

The RIN_1 input case also reflects this improvement with the intersect of the output not on the scale and the difference in the outputs at an input RIN_1 of 120dB/Hz up from 4dB previously, to 6.5dB.

From this it can be reasoned that if a coupling ratio of 0.1 for the error signal is used, with the correct amplitude adjustment implemented, a reduction in noise of over 6dB may be achieved. Whilst this is a significant reduction in noise, this is not as great a reduction as can be achieved in fibre system [79]. This is most likely due to the extra photodiode and amplification components that are used in the free space system.

When encompassing the coupling ratio of 0.1 into the main advanced Volterra program to check the effect upon IMD_3 compensation, very little change was seen. A very small reduction of approximately 2dB in the level of compensation was observed.

With a 6dB reduction in noise, combined with a minimal effect on the IMD_3 compensation, there is a strong claim for using a coupling ratio of around 0.1.

4.3 Summary and conclusions

In this chapter the noise performance of the system has been simulated the effect of the feedforward compensation technique upon this performance investigated.

The theory suggests a small drop in output noise for the system due to the employment of feedforward compensation, providing the noise contribution from the primary path laser diode, that of RIN_1, is significant, as it is only this noise source that is compensated for.

When a coupling ratio for the main path and error path signals of 9:1 was employed, further reductions in the output noise were observed, indicating that this should be considered practically. However, practically this is harder to achieve in the simulations, as will be explained in greater detail in the following chapter and hence will remain as a simulated result only.

Chapter 5 – Practical Work

In this chapter the experimental arrangement and result of the optical free space feedforward non-linear correction system are shown. Also described are the important electrical and optical characteristics, the problems encountered and the subsequent solutions.

As stated in section 3.5, the practical set-up of the system varies somewhat from the theoretical ideal. Firstly, the system was realised in two stages, with the first stage having an electrical correction path, and the second stage being fully optical. The reasoning for this was that it allowed the testing of the principle without its full complexity, thus breaking down the full system realisation into smaller, more testable, steps.

5.1 Stage one set-up, with electrical correction path

This initial system included one laser diode, two photo diode receiver-amplifiers, a pellicle beam splitter, three electrical combiner/splitters, two additional wide band amplifiers, two optical antennas and an attenuator. Adapted from figure 2.10, the first stage system diagram can be seen in figure 5.1 below.

The laser diode was a single mode, Fabry Perot laser, with a transmission wavelength of 850nm. It was built into a module (PMA LD module) with an analogue modulation circuit capable of modulation –3dB@25MHz, supplied by Laser 2000 UK Ltd. The LD was powered to threshold by a 5Vdc supply, with a separate voltage, of a maximum of 1Vdc supplying the bias, (which when applied, resulted in a continuous wave output).

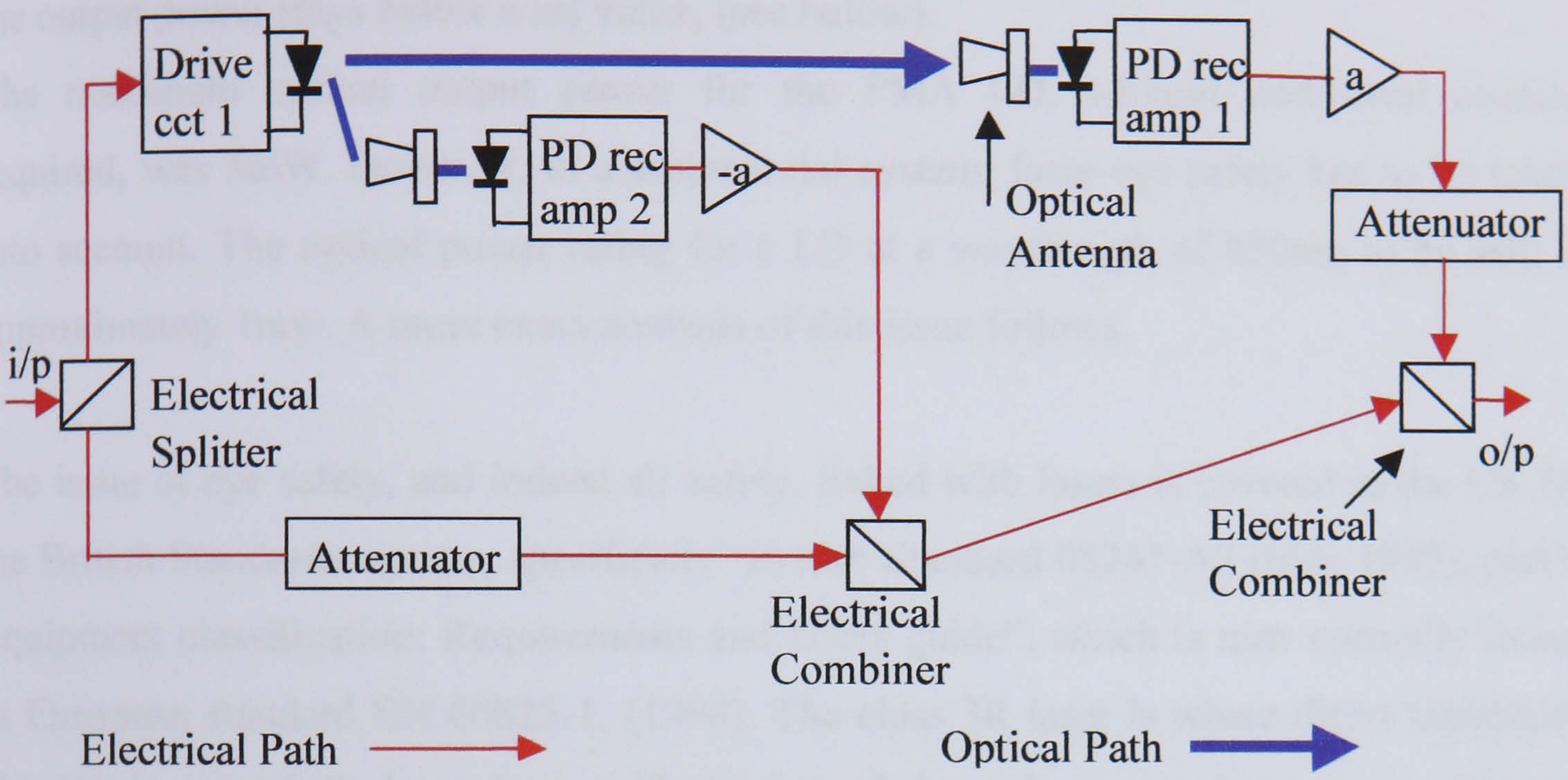


Figure 5.1 *First stage of practical system realisation, with electrical correction path*

To modulate, the bias supply was reduced slightly and combined with the analogue signal, with the peak current value of the analogue signal representing I_{peak} in equation (3.9) Chapter 3. The supply from the two signal generators was a combined 3dBm signal at 20MHz and 20.5MHz. This translated as an input voltage, into a 50Ω matched system, of approximately $0.3V_{peak}$. Although the information on how the PMA LD converted the supplying voltages into the threshold, bias and modulating currents was not available, if I_{bias} and I_{peak} of equation (3.9) Chapter 3, are analogous to V_{bias} and V_{peak}, by a common resistance, it was found that, by setting the input voltage to 0.7Vdc, an approximate modulation index of around 0.4. was achieved. For this purpose, a small circuit was designed and built. Included in the same unit were the 5Vdc threshold supply and the 5Vdc to –5Vdc supply for the op-amps used within the system, (see later). A photo of this unit can be seen in figure 5.4 at the end of this section. The circuit not only combined the dc bias voltage with 20-20.5 MHz modulation supply, but also offered circuit protection via a set of six diodes. A potential divider was added which gave greater bias voltage control on the power supply unit. The PMA LD has built in focusing optics that can be

adjusted by the use of a supplied key, and required no additional cooling provided that the output power stays below a set value, (see below).

The maximum optical output power for the PMA LD, without additional cooling required, was 5mW. However, in a commercial system, laser-eye safety has to be taken into account. The optical power rating for a LD at a wavelength of 850nm to be safe is approximately 1mw. A more exact analysis of this issue follows.

The issue of eye safety, and indeed all safety, linked with lasers is covered in the UK by the British Standards agency, specifically "British Standard 08241-A3 (May 1997), part1: Equipment classification: Requirements and Users guide", which is now normally found as European standard EN 60825-1, (1994). The class 3R laser is where direct intrabeam viewing is potentially hazardous, with the hazard dependant upon laser power, time of exposure and area of exposure, i.e. the energy density. Classes 1, 1M, 2, 2M and 3R are all eye safe providing certain rules, such as those mentioned above, are adhered to. Classes 3B and (especially) 4 are generally considered hazardous, with strict precautions in use required. For varying exposure times, a number of equations are listed to give limits of power and power density. Class 3R is similar to class 2M, and deals with lasers in the IR spectrum, i.e. above 700nm.

For exposure from 10s up to $3x10^4$s (or 8hrs 20mins, which would be a very long time to be exposed directly to a beam) the equations for permitted power are as follows:

$2 \times 10^{-3} C_4 C_7$ W for $\alpha \leq 1.5$mrad **(5.1)**

where $C_4 = 10^{0.002(\lambda - 700nm)}$ and $C_7 = 1$ for $700nm < \lambda < 1150nm$ **(5.2)**

and

$3.5 \times 10^{-3} C_4 C_6 C_7 T_2^{-0.25}$ W for $\alpha > 1.5$mrad **(5.3)**

where $C_6 = 1$ for $\alpha \leq \alpha_{min}$

$= \alpha/\alpha_{min}$ for $\alpha_{min} < \alpha \leq \alpha_{max}$

$= \alpha_{max}/\alpha_{min}$ for $\alpha > \alpha_{max}$

where $\alpha_{min} = 11$mrad and $\alpha_{max} = 0.1$rad, (at 100mm from source) **(5.4)**

and $T_2 = 10$ for $0 < \alpha \leq 1.5$

$$= 10 \times 10^{[(\alpha\text{-}1.5\text{mrad})/98.5]} \qquad \text{for } 1.5 < \alpha \leq 100$$

$$= 100 \qquad \text{for } 100 < \alpha \qquad \textbf{(5.5)}$$

where λ is the laser diode wavelength and α is the angular subtense of the source.

The power limit from equation (5.1) for λ = 850nm is then 4mW. This is for α <1.5mrad or 0.86^0. For a larger angle, i.e. α = 10mrad (or 5.73^0), equation (5.3) is used with the power limit being 1.87mW.

The equivalent limits for a class 1 laser, which are considered safe under standard operating conditions using the class 1 equivalents of equations (5.1) and (5.3), are 0.78mW and 0.37mW respectively.

Of course, it is not the optical power that necessarily does any harm, but the optical power density, in W/m^2. For example, a low power device with a very small spot size can have the same power density as a high power device with a large beam spread. The aperture diameter for the measurement of laser irradiance and radiant exposure for the eye, for wavelengths of 400nm to 1400nm is 7mm, giving an area of around $38.5mm^2$. Therefore, for smaller spot sizes smaller than this, care must be taken. The maximum power density for skin exposure for a time over 10s is $4mW/mm^2$.

With this system being for proof of principle, and as the experimental system is set up under controlled conditions, the issue of laser-eye safety does not have to be adhered to strictly. However it is advisable to set the LD power levels at approximately the levels mentioned above. Therefore LD1 was set at 2.5mW, after the focusing optics, although, as the spot size is smaller than that stated above, care must be taken. This allows for a good signal to noise condition, and offsets some of the losses in optical power attributed to the beam splitter. The beam splitter is a pellicle device supplied by Edmund Optics Ltd. Pellicles are very thin nitro-cellulose membranes and have almost imperceptible absorption. When the incident beam hits the beam splitter at 45^0 there is a 40% transference along the line of the incident beam and a 40% reflectance at 90^0 to the incident beam, giving approximately 1mW per beam. Compared to a standard beam splitter, there is no lateral transference in either beam that can lead to a ghost image, as

the second surface reflection is super imposed on the first surface reflection. The principle is further explained in figure 5.2 below.

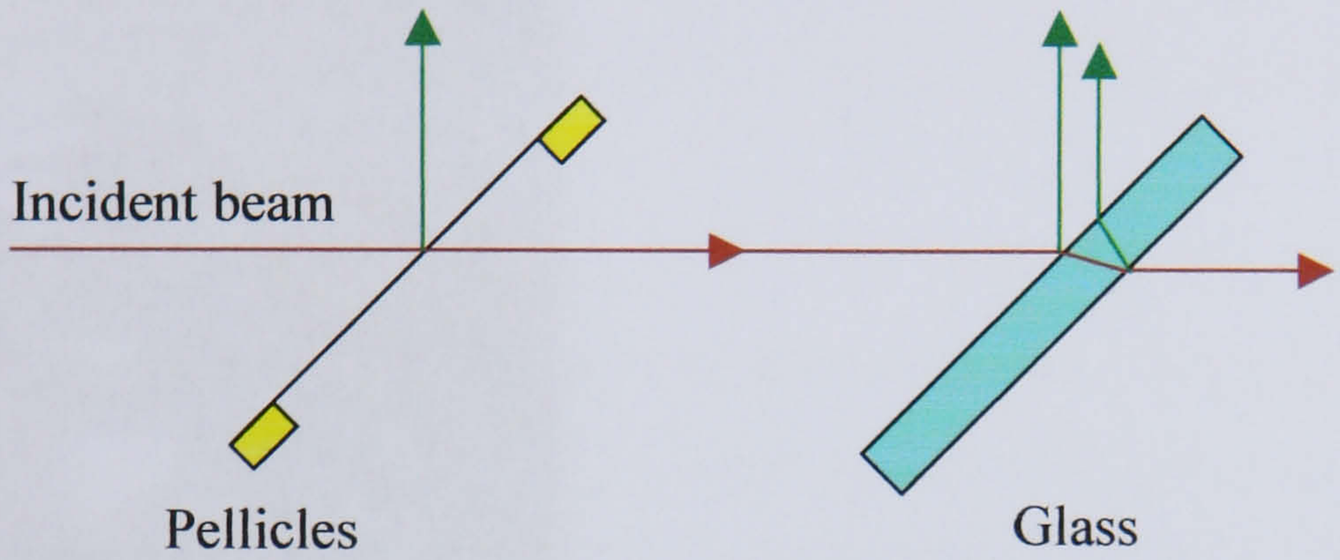


Figure 5.2 *Illustration of a pellicle beam splitter compared to conventional beam splitter*

Before arriving at the photo detectors, (PDs), the signal from the LD encounters the optical antenna.

Developed and patented by colleagues at Warwick University, the optical antenna improves the collection efficiency of a receiver by transforming light rays incident over a large area into a set of rays that emerge from a smaller area. This implies that smaller photodetectors can be used, which decreases the capacitance, the cost, and improves receiver sensitivity.

The use of omnidirectional and directed concentrators, used in conjunction with multilayer bandpass filters, effectively rejects unwanted out-of-band radiation and increases the effective area of the photodetector.

Compared to hemispherical concentrators, the Dielectric Totally Internally Reflecting Concentrator, or DTIRC, offers a higher concentration and the possibility of using flat thin-film optical filters [80].

Compared with Compound Parabolic Concentrators (CPCs), the DTIRC has two advantages: a smaller size (its size can be nearly 1/5 of that of a single dielectric CPC if the front surface arc angle is large enough), and a higher concentration. This is because of

the use of a curved front surface at the entrance of the concentrator, and the large refractive index of the dielectric materials. Figure 5.3 shows a photograph of a DTIRC.

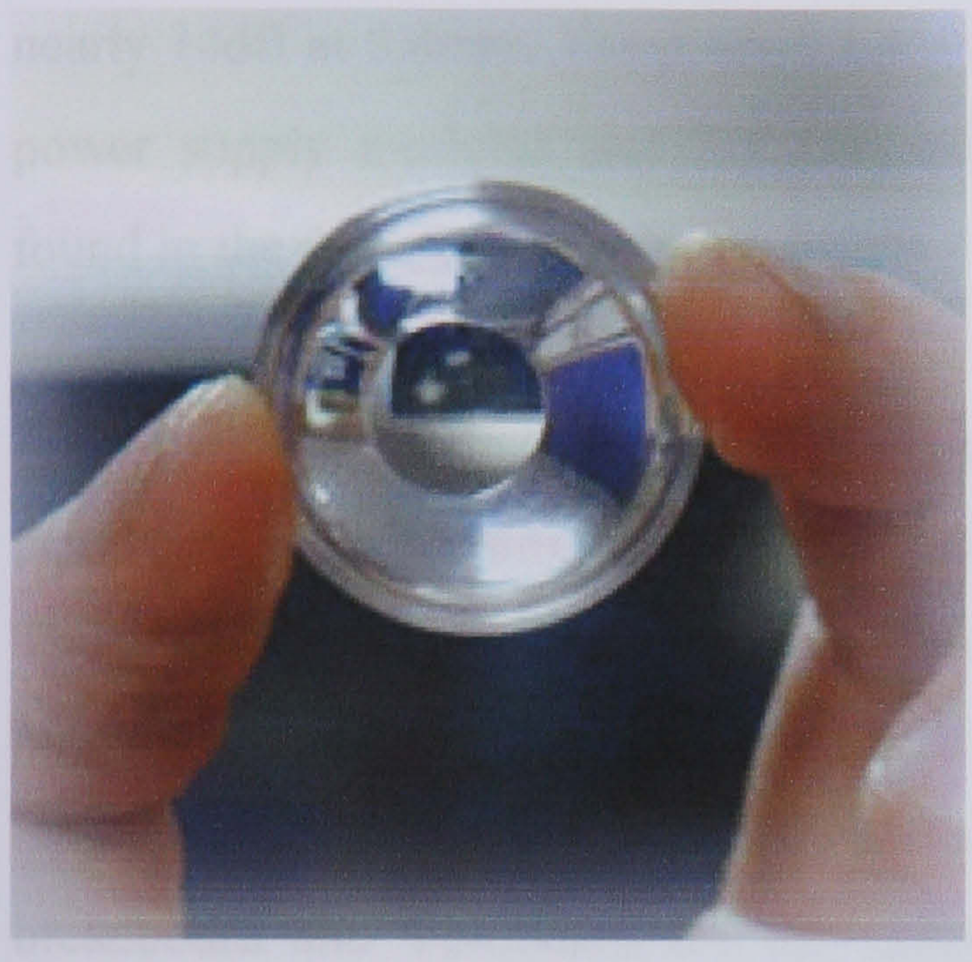


Figure 5.3 *DTIRC photograph (courtesy of Roberto Remirez-Inguez, University of Warwick)*

Further in depth analysis of the DTIRC can be found in [81].
(The optical antennas fabricated at Warwick University use high-performance optical material developed by Nippon Zeon Co., Ltd.: Zeonex 480R and Zeonor 1420R).

The receiving surface area the PDs used in this system are very small (approximately 1mm in diameter), making the appropriate alignment and focusing of the beam from the LD a difficult and time consuming process. By fastening a DTIRC to the front end of each PD, the receiving surface area is effectively greatly increased, i.e. diameter 5mm, therefore an approximate 25 times increase in surface area. This makes alignment of the LD beam a lot simpler and gives a greater range of focusing along the axis of the LD beam. Whereas the PD could only be moved around 2cm along the axis and remain within a particular focus length, this was increased to around 6cm with the introduction of the DTIRC. It was secured in this instance to the PD using Sellotape, with a neutral

refractive index gel placed in the contact area to ensure a seamless transmission of the optical signal to the PD.

The Electro-Optics Technology Inc ET-2030A PD units have a built in wide band amplifier with a frequency range from 75KHz to1.2GHz and an amplification of 25, or nearly 14dB at 830nm. These were supplied by Laser 2000 UK Ltd. They have their own power supply modules and 50Ω matched BNC outputs. (Further information can be found in the data sheets in Appendix D).

From both PDs, the signal was transmitted to additional wideband amplifiers, the main path being non-inverting amplifier and the error path having an inverting amplifier.

The amplifiers were designed and built at Warwick University on PCB, using a surface mount current feedback twin op-amp IC from Dallas-Maxim (MAX4117), with all the other components used also of surface mount technology. At these frequencies breadboard and Veroboard are too capacitive, hence a PCB with constant impedance microstrip techniques were employed, i.e. short rounded tracks rather than long tracks with 90^0 corners. Surface mount components give better high frequency performance than through-hole components, due to having shorter leads and hence lower parasitic reactances.

With the peak output from the PDs at around –25dBm, a gain of around 10 to 15dB was desirable, so as to increase the signal close to that of the reference path, with the rest of the signal difference being cancelled by an attenuator. This gain was obtained in two stages for each amplifier, thus keeping any amplifier distortion to minimal levels by taking advantage of the twin amp IC technology..

The non-inverting amplifier actually consisted of two inverting sages back to back, while the inverting amplifier had an inverting stage followed by a non-inverting stage. The data sheets for the MAX4117 op-amp along with the three full circuit diagrams of the amplifiers, complete with power line capacitive decoupling, seen in figures D2 and D3, can be found in Appendix D.

With the manufacturers recommendation for the feedback resistance $R_{f/b}$ to be set at 680Ω, R_{in} was set at around 200Ω to give a gain of around 10 to 12dB for the inverting and non-inverting amplifiers. In addition to above circuitry, a 24Ω resister was placed in

series on the output of each amplifier to isolate the current feedback path from any capacitance on the coaxial cables. This helped stability by reducing the possibility of oscillation. The outputs of the amplifiers were also a.c. coupled to prevent any unwanted d.c. offsets affecting the cancellation process. It was found in the simulations of Chapter 3 that any offset leads to a serious reduction in IMD cancellation.

Practically, both amplifier performed as expected achieving the required gains, whilst introducing very little noise into the system above that already observed.

Linking the whole system together were three electrical combiners/splitters, supplied by Mini Circuits. These are 2 way-0^0 (i.e. no phase shift) devices, with an insertion loss at 20MHz of 3.3dB. Further information can be found in the data sheets in Appendix D.

In the reference path an attenuator was placed so the two correction path signals could be amplitude matched before combination. The attenuator was calibrated and could attenuate a signal from 1dB to 100dB. Placed in front of this attenuator was a zero gain buffer amplifier to isolate the correction path circuitry from the signal generators and a.c. couple the system, as for previous reasons. Another attenuator was placed on the main path for the same reasons as above.

Also used in this practical arrangement were two signal generators, a spectrum analyser, a digital storage oscilloscope and a pair of twin output power supplies. The two signal generators supplying the fundamentals were a Rohde & Schartz SML03 (9KHz...3.3GHz) and a Marconi Instruments 2019 (80KHz...1040MHz). These were set to generate signals at 20.5MHz and 20MHz respectively, with an amplitude of -1dBm. The spectrum analyser used was an Advantest R3131 (9KHz...3GHz), and the oscilloscope a Tektronix 2465A (350MHz). As previously stated, in the interests of tidiness, the power supplies for the LDs and op-amps were incorporated into a single unit, as shown in figure 5.4 below.

The whole system was matched for maximum power transfer to 50Ω and to reduce reflections, though at the relatively low frequencies being used, it was not vital for this to be so, the subject of delay being dealt with in the following section.

All the optical components were post mounted, with the laser diode(s) held via v-block mounts, which were all supplied by Elliot Scientific Ltd and Edmund Optics Ltd. The

various post mountings were held on a Δ-rails. A picture of the final set-up can be seen in figures 5.6 and 5.7 in the next section.

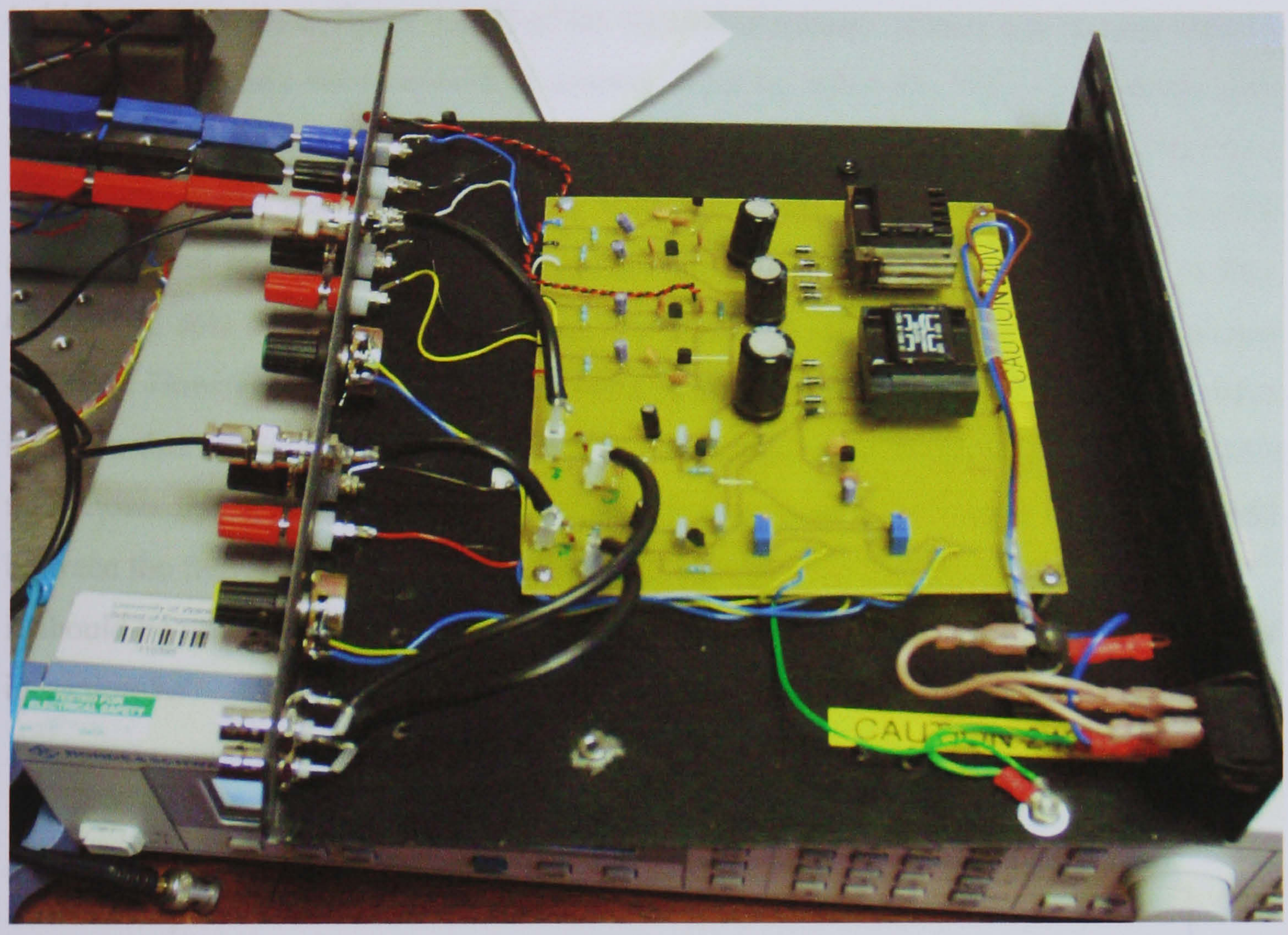


Figure 5.4 *Power supply and modulation supply unit for the laser diodes and op-amps*

5.2 Stage one results

The 20MHz/20.5MHz input signal was fed into the system described in the previous section, with figure 5.5(a) showing the non-corrected output of the free space communications link. It can be seen that the fundamental inputs were at a level of –22dBm, and the IMD_3 was approximately –45dBm, only 23dB below the fundamental. Significant IMD_5 levels were also clearly seen. This output signal was subject to all the

combiners, and their respective losses throughout the system, and hence the output level was relatively low. The noise floor was around –70dBm.

As shown in Chapter 3, the input signal was combined with the appropriately amplitude and phase adjusted reference signal, to produce the error signal seen in figure 5.5(b), (which corresponds to figure 3.1(b) of the simulated results). Here it can be seen that the fundamental signals were reduced to around –42dBm, while the IMD components were essentially unaffected. This error signal was the same for both stage one and stage two. This error signal was then combined with the main path signal, via some further amplitude and phase adjustment, to produce the corrected signal seen in figure 5.5(c). Comparing this figure to figure 5.5(a), the fundamentals were unaffected, at around –22dBm. However, the IMD products reduced by over 20dB. Though this is somewhat less than that of the 50dB drop predicted in Chapter 3, it is a sufficient reduction in IMD that shows the applicability of the method. It also gave a difference of around 45dB between the fundamental signals and the IMD_3 signals.

It should be noted that the noise characteristics changed, and these will be discussed later in section 5.7

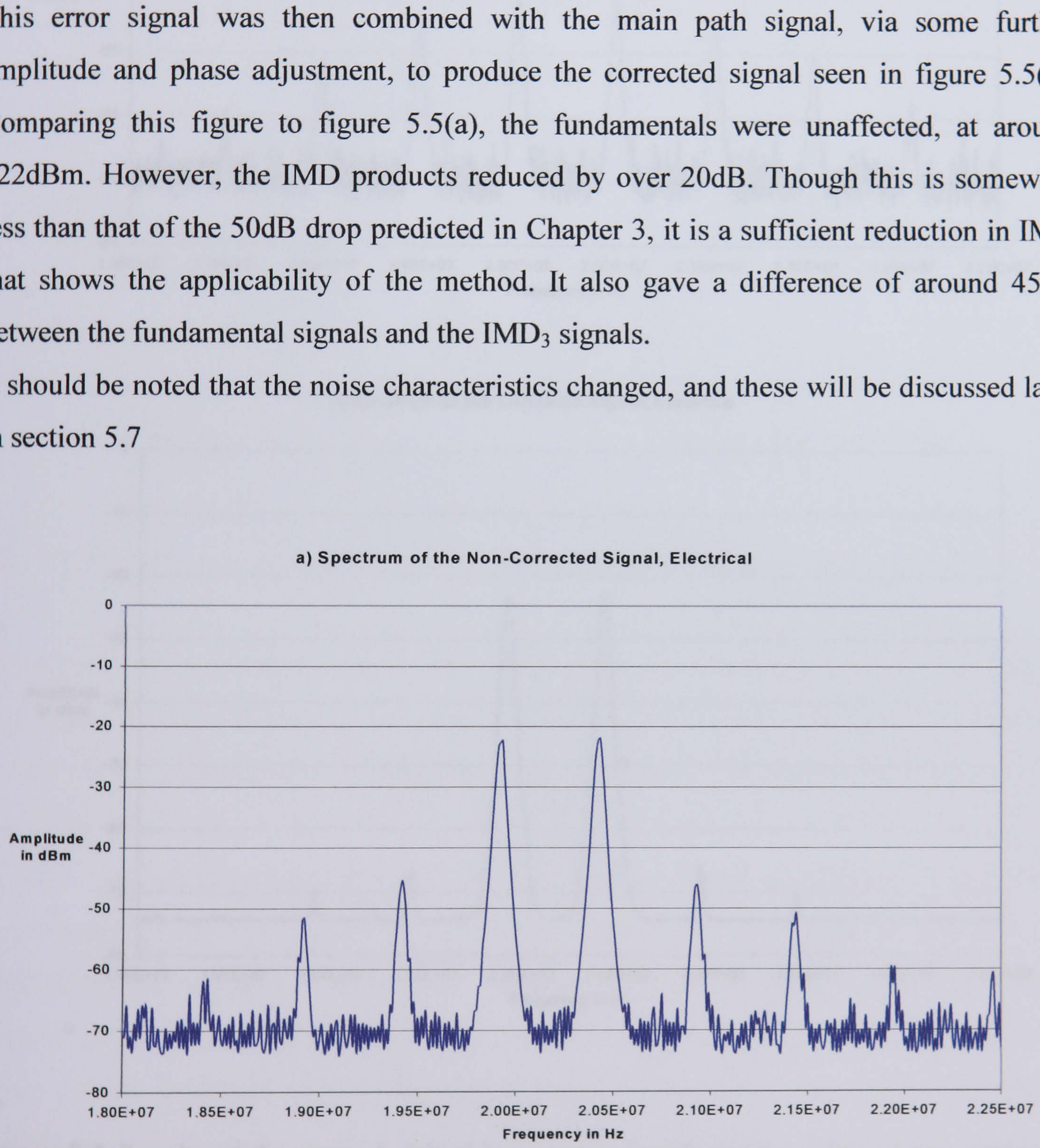


Figure 5.5 Results of the stage 1 cancellation when using Agilent, with (a) Non-corrected signal, (b) error signal and (c) corrected signal

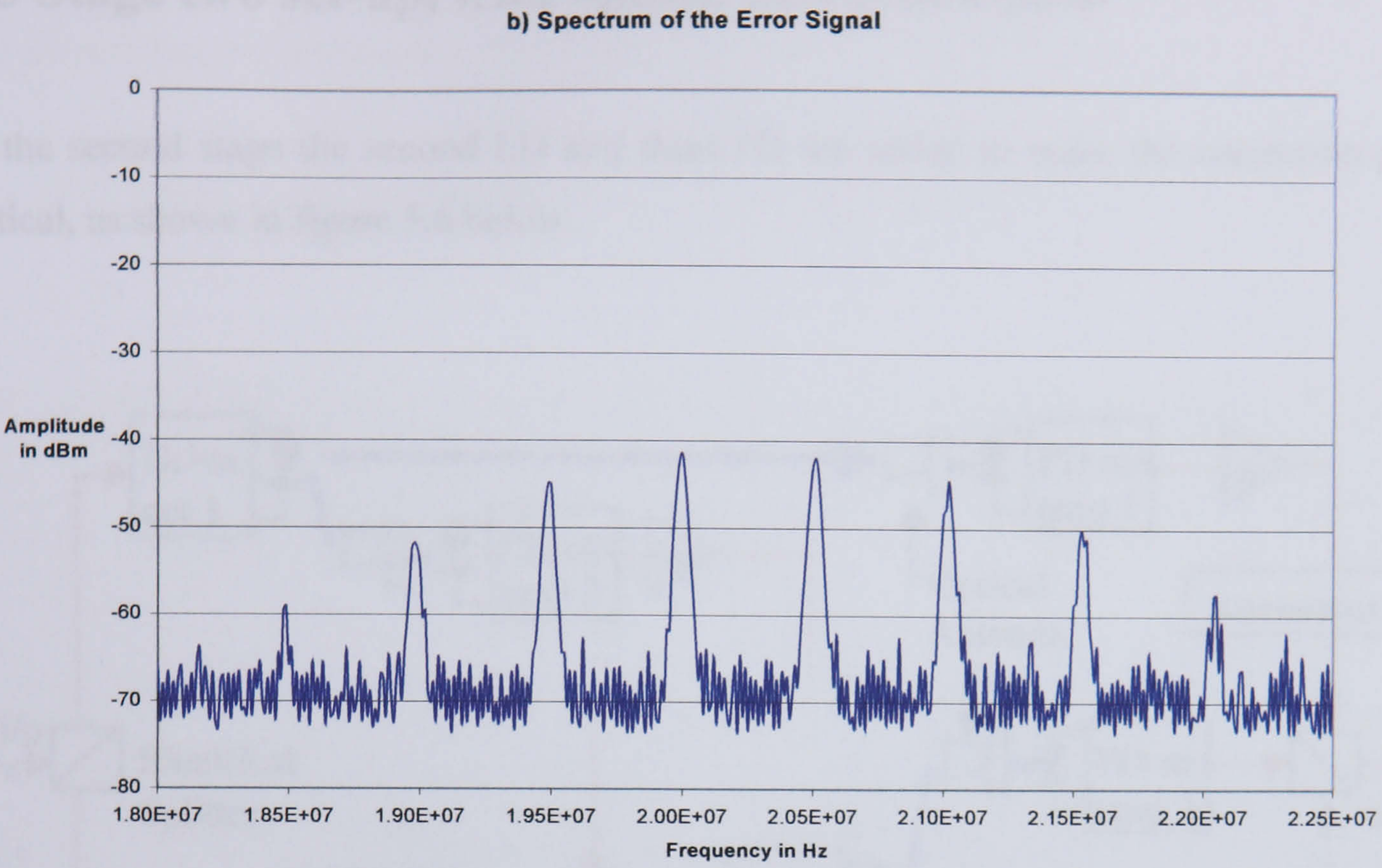


Figure 5.5 *Results of the stage 1, (electrical error signal) system, with; a) pre-corrected signal, b) error signal and c) corrected signal*

5.3 Stage two set-up, with optical correction path

In the second stage the second LD and third PD are added to make the correction path optical, as shown in figure 5.6 below.

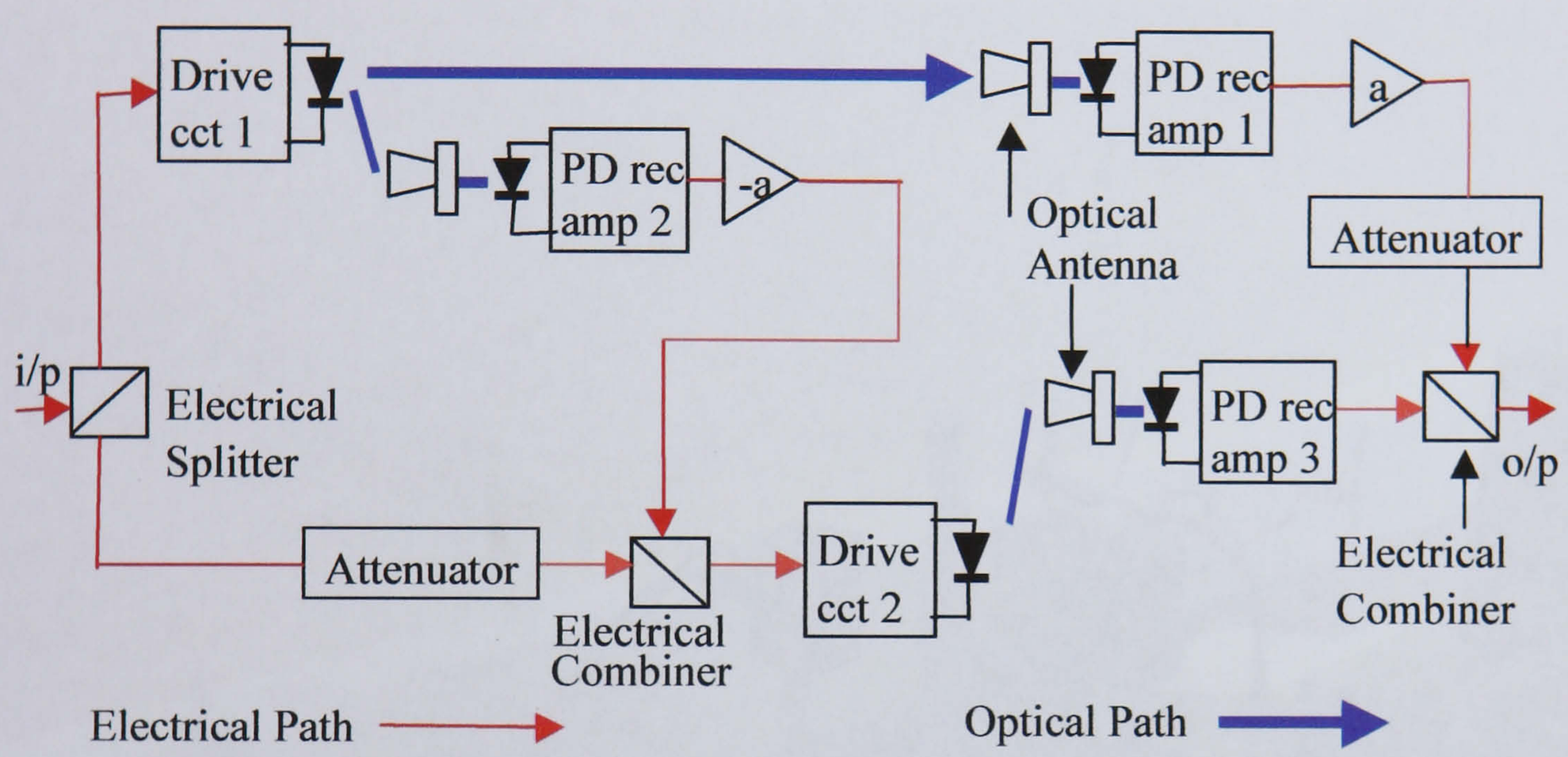


Fig. 5.6 *Second stage of practical system realisation, with optical correction path*

As can be seen in the above figure, the optical correction path is in place and the system now resembles that of the one introduced in figure 2.10. Though the delay loops are not shown above, some delay to adjust any phase difference was required in stage one and also stage two.

In stage one, observations revealed that the reference signal needed to be delayed by approximately 6 to 7ns, (due to the signal appearing blurred due to the distortion, it was difficult to be any more accurate than this). A simple experiment with the standard coaxial cables used in the system revealed that a length of 120cm gave a delay of 22ns. Thus, a cable length of around 36cm was inserted into the reference path. This proved sufficient.

In the second stage, with the additional LD/receiver unit inserted for fully optical operation, an additional cable of 36cm was required on the main path to provide

sufficient delay for the maximum cancellation. Again, the second LD's threshold, bias and modulation signals are supplied by the unit shown in figure 5.4.

The final system set-up can be seen below in figures 5.7 and 5.8.

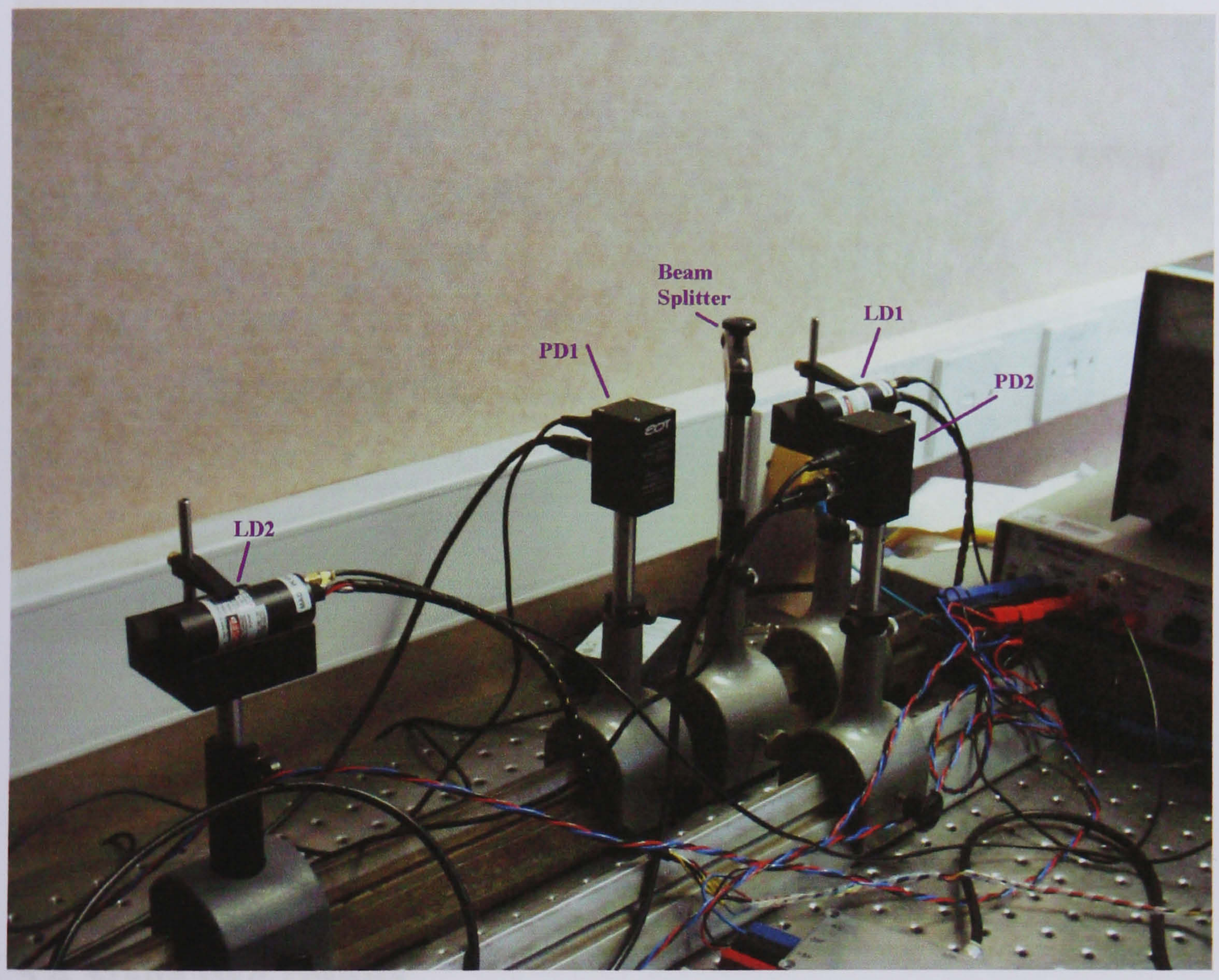


Figure 5.7 *Stage one optical components*

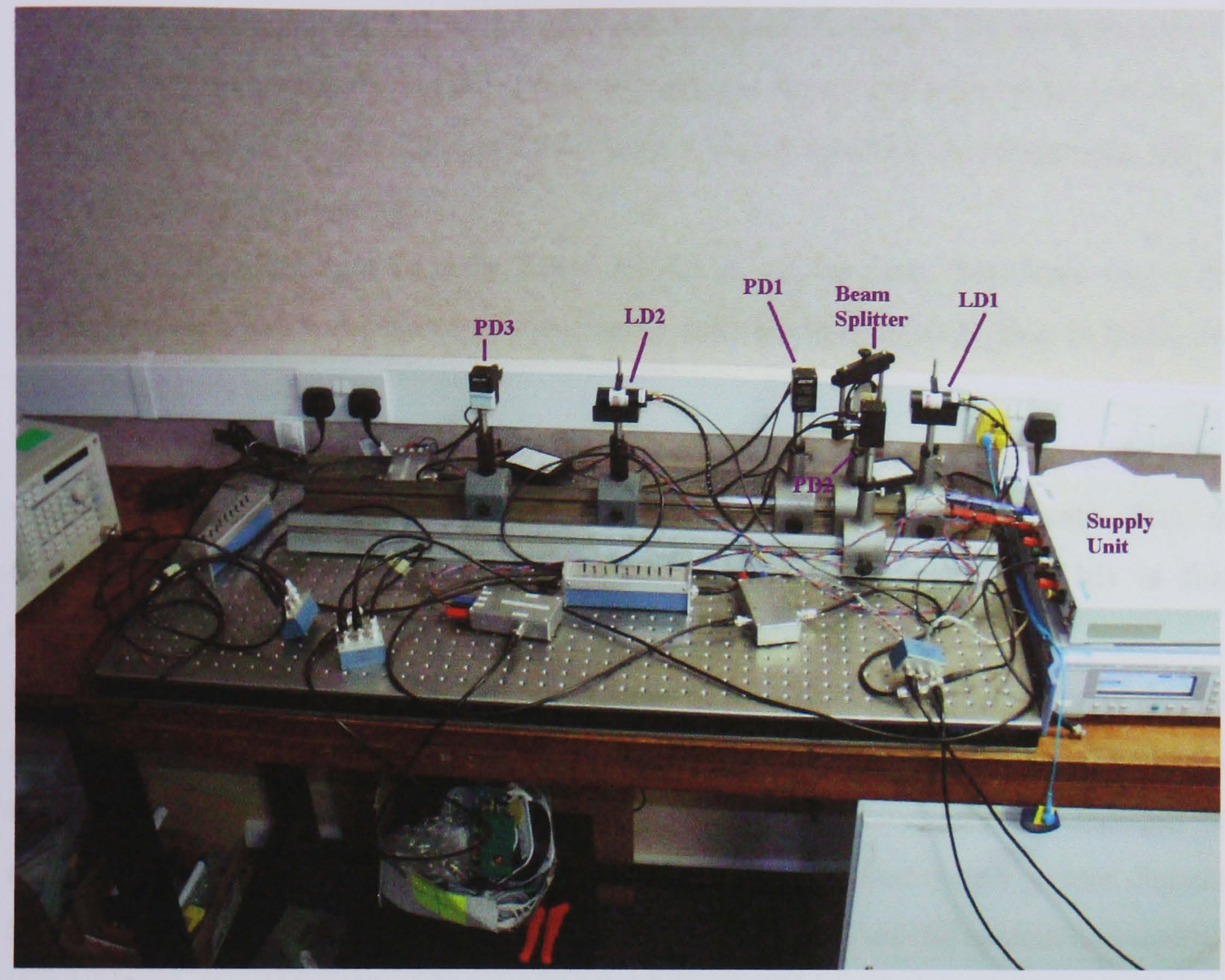


Figure 5.8 *Final system set-up*

5.4 Stage two results

Following on from stage one are the results for the fully optical system highlighted in the previous section of 5.3. Again, the 20MHz/20.5MHz input signal was fed into the system, with figure 5.9 showing the pre-corrected signal and the post-corrected in 5.9(a) and 5.9(b) respectively. The error signal is not shown here, as it is the same as in the previous section.

The only difference between the non-corrected signal of the two sections was that the whole signal, i.e. the fundamentals, IMDs and noise floor, were at a level of around 10dB

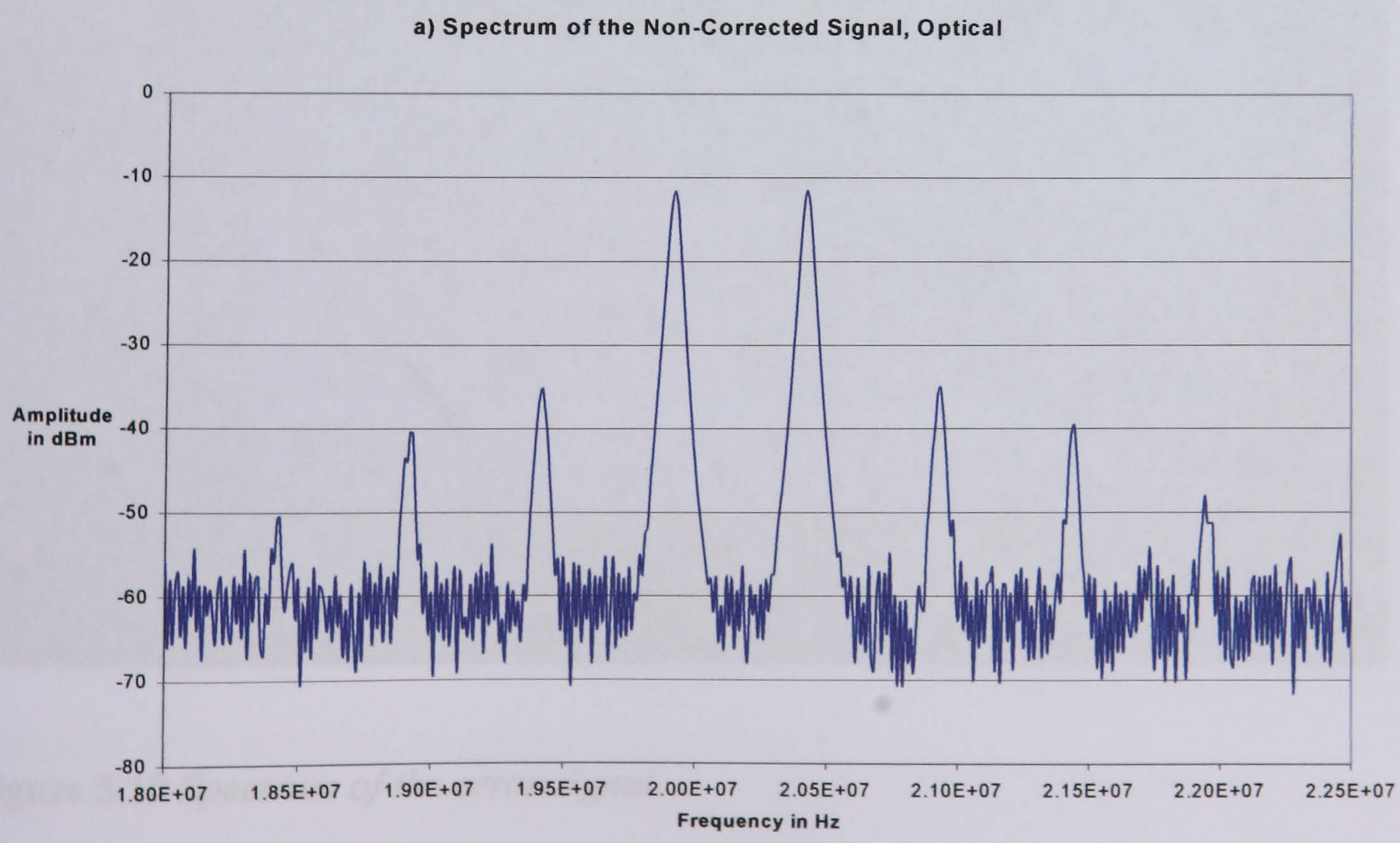
higher in the optical non-corrected signal case. This was due to the difference in amplitude adjustment needed in the fully optical system's case. The difference of 23dB between the fundamentals and the IMD_3 was still the same. The noise floor was at around –60dBm, though it would also appear to have a higher level of deviation from the mean than in the stage one system.

Looking at the corrected case in figure 6.2(b) it can be seen that there was a 20dB reduction in the IMD_3 on the higher frequency side, but only a 10dB drop in IMD_3 on the lower frequency side.

When the phase compensation coaxial cable was altered to attempt to rectify this, the reverse IMD characteristics were observed. This seemed to indicate that more sophisticated delay techniques for phase adjustment were required. This is further discussed in the following chapter.

From figure 5.9 it may be seen that an average of a 15dB drop in IMD is associated with a minimum reduction of 10dB in IMD_3. This is still a useful reduction in IMD and shows that the system works at a fully optical level and that, with the appropriate changes in the amplitude and phase adjustment, this technique could be applied over a greater distances.

Shown in figures 5.10 to 5.12 are pictures of the spectrums of the optical non-corrected, error and corrected signals, corresponding to those in figures 5.5(b) and 5.9.



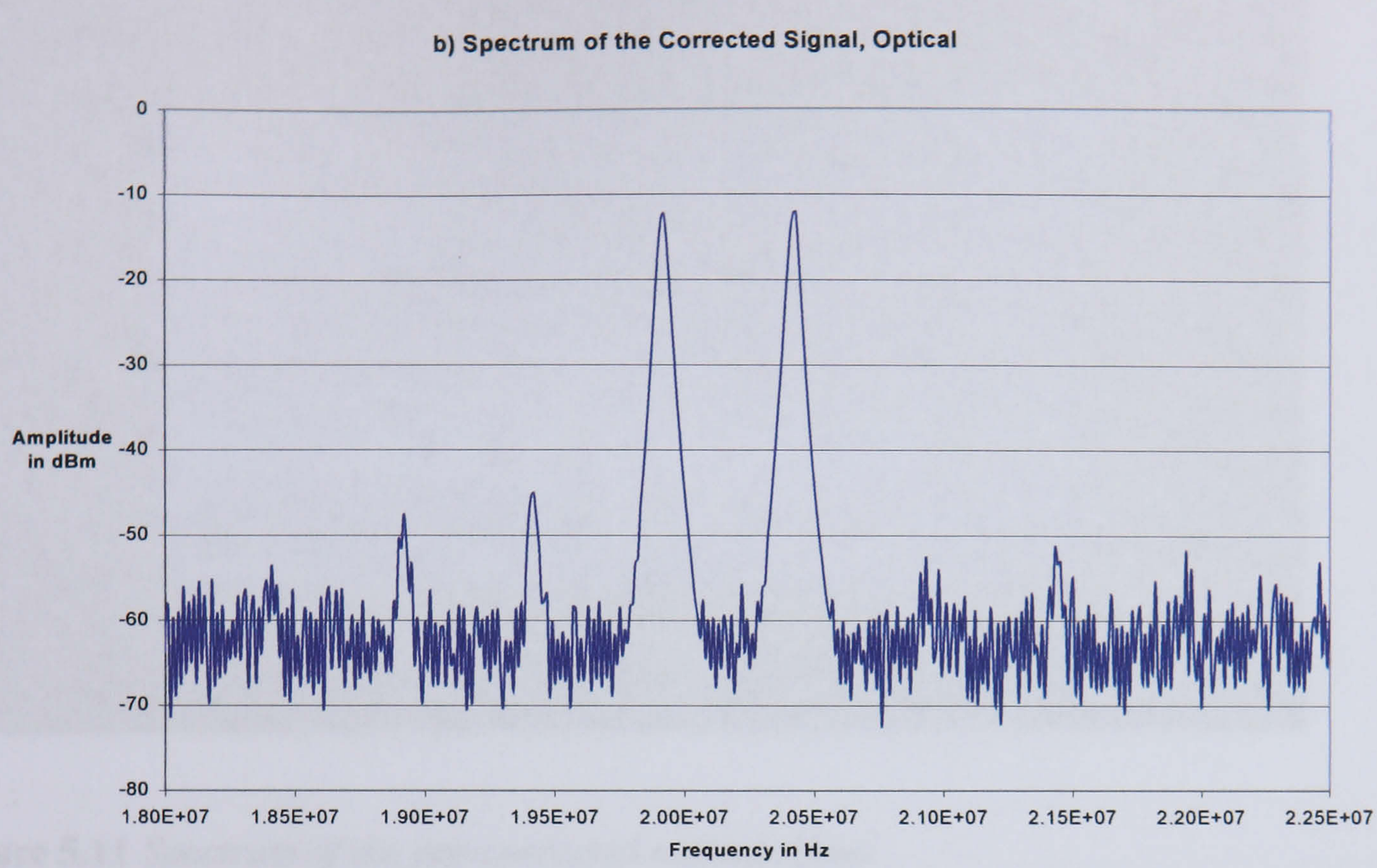


Figure 5.9 *Results of the stage 2, (optical error signal) system, with; a) pre-corrected signal and b) corrected signal*

Figure 5.10 *Spectrum of the error signal*

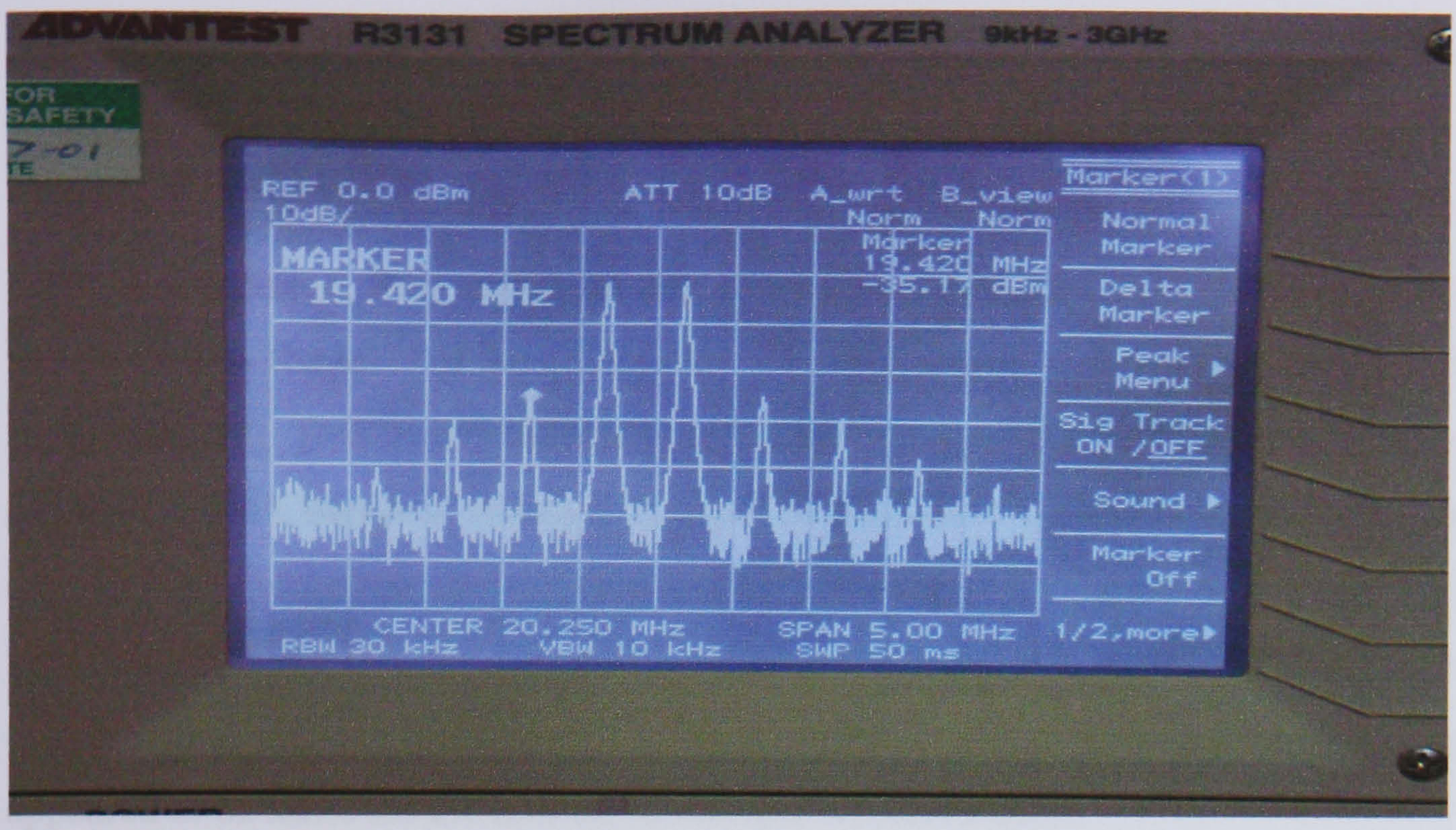


Figure 5.11 *Spectrum of the pre-corrected optical signal*

Figure 5.12 *Spectrum of the corrected optical signal*

5.5 Noise analysis

Looking back at figures 6.1(a) and 6.1(c), it appears that the noise is reduced in both mean and variance. This could be that the RIN of the main path laser diode has been reduced below the level of the next dominant noise source, that is, at the level of –75dBm in figure 6.1(c). This would indicate a reduction in noise of over 5dB which, while much as predicted in Chapter 4, is better in terms of equivalent performance than the reduction seen in IMD. Moreover, for the fully optical case shown in figures 6.2(a) and 6.2(b), the reduction in noise is a far more realistic at about 2dB. When modulation signals were disconnected, and the bias current increased slightly to compensate, the noise observed appeared to be same for the non-corrected and corrected case.

Therefore, in conclusion, it seems that whilst the above results indicate that some noise reduction in this system is possible, it needs further work, with more accurate equipment to fully gauge the extent of the possible main path laser RIN noise reduction.

This is further discussed in the following chapter.

5.6 Conclusion

To conclude this results section, useful reductions in IMD have been practically demonstrated in both the stage 1 electrical error path system and, more significantly, in the stage 2 optical, error path system.

This justifies and confirms the previous simulation work of Chapter 3. There clearly improvements that can be made in respect of the practical system's performance. These improvements are discussed in the following chapter.

Chapter 6 – Conclusions and further work

In this chapter, the conclusions of the work done in this thesis, as well as a summary of further research, and the improvements that could be made to the system, will be detailed.

In Chapters 3 to 5, the simulations, practical set-up and results of the operational optical freespace feedforward non-linear correction system were presented and discussed.
As previous work had shown, [40], [42], [46] the non-linearities created by a laser diode when modulation with an RF signal, can be reduced be a number of different methods. In this work, one of these methods, namely feedforward correction, was adapted to a free space format. This adaptation took place first in software simulations, then in a practical hardware laboratory experiment.

The theoretical simulations required the use of an advanced mathematical technique, known as the Volterra analysis, to accurately model the system. This was due to the high frequencies of the RF signal used, that of a UMTS carrier, around 2GHz. For frequencies over approximately 100MHz the laser diode's non-linear characteristics can not be modelled by a Taylor's series and require the more sophisticated technique.

The simulations in Chapter 3 (section 3.4) show a reduction in IMD_3 of around 40dB, well in line with previous research using this technique over optical fibre [46], and a very significant reduction. Chapter 4 also showed that could also be some reduction in RIN noise expected.
Chapter 5 showed the practical set-up and the results of the practical tests. This was done in two stages, an electrical-optical set-up and a fully optical set-up. Both revealed significant reductions in IMD_3, with a final test reduction of approximately 15dB. Whilst not as good (for a number of reasons discussed within Chapter 5) as the theoretical simulations of 40dB, a 15dB reduction in third order intermodulation interference is very promising, however there are several steps that could be taken to improve this figure.

It would be beneficial to investigate the effect of varying coupling ratios, as illustrated in Chapter 4. This would require far more sophisticated electrical combiners than those currently employed, combined with variable gain amplifiers so as to achieve the correct relative signal amplitudes. The work in Chapter 4 does suggest that some benefit in noise performance is possible, (figure 4.5).

In a commercial system DFB lasers would have to be used, as Fabry Perot LDs are not generally suitable for modulation frequencies of up to 2GHz plus. Also, as they operate at a wavelength of around 1310nm and 1550nm, they can be driven at a higher power whilst still being 'eye safe' when compared to LDs transmitting at around 800nm. This results in a more favourable noise performance, as the dominant noise source is more likely to be the RIN of the main path laser, which, as Chapter 4 demonstrated, can be compensated to some degree, (see figure 4.4), especially if selective coupling ratios are employed.

The major drawback of DFB LDs is financial, as their cost is considerable when compared to FP LDs. Added to this is the fact that silicon photodiodes are unsuitable at the longer transmission wavelengths (e.g. 1300nm and 1550nm) and gallium-arsenide detectors are then needed. The cost of these is also considerable (albeit reducing as time goes on) when compared to those used in this study.

It has been stated on several occasions in this work that the ultimate limiting factor to the amount of IMD cancellation possible is the IMD introduced by the correction path laser. Pre-distortion for this laser can reduce the IMD by around 10dB (section 2.5.2), so that this limit can be pushed further. Nevertheless, the drawbacks of pre-distortion, such as the lack of ability of self-adjustment, will remain. However, these drawbacks are not as critical in this instance as they will only affect the correction path LD. Pre-distortion is also a relatively simple and inexpensive technique to implement.

The issue of beam alignment can be a problem. For example, if a system is placed on the two roof tops of two tower blocks, building sway due to wind could cause major throughput problems. The transmitters and receivers could be mounted on sophisticated

platforms that keep the systems stable. However these are likely be costly due to the precession needed to maintain the required beam spot position.

A better option would be to spread the beam from the laser diodes so that the receivers are always within the beam spot. The use of the optical antenna on the receiver, as shown in Chapter 5, would help this. The optical output power of the LDs would need to be increased, but as long as the power density stays within eye safe levels, this should not be a concern. The only problem is that, the beams should not be spread to the point that they interfere with each other, or any other co-located systems, as this could lead to signal degradation as well as raise some issues of signal security. A blocking back panel at the receiver site preventing signal overshoot would be one necessary step, see figure 6.1 below.

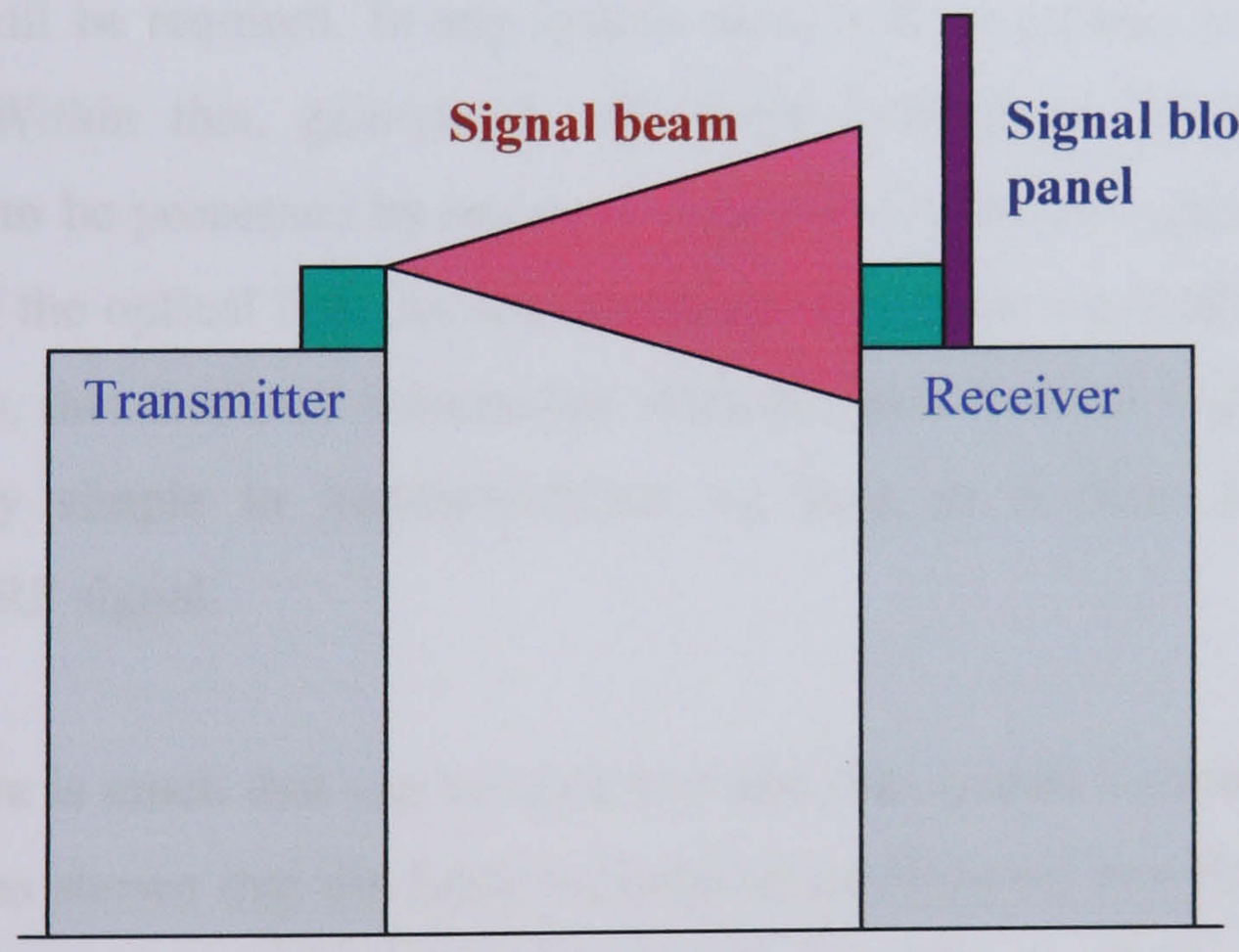


Fig. 6.1 *Example of a commercial roof-top system, with signal blocking panel*

Any commercial system would require automatic gain and phase adjustment to respond to any variations within the system to ensure the optimum cancellation of non-linearities. The most likely of these are amplitude and phase variations due to changing laser diode

characteristics that occur with ageing, though the feedforward technique does partially accommodate for this.

The main parameters that need to by considered for the control unit are; the dynamic range over which the amplitude and phase can be varied, the frequency response, and the size. The gain and phase adjustments can be undertaken via the amplifiers and delays already within the system, (figure 2.10), or by separate additional units, placed in series with the error path amplifiers, rather than in series with the delay lines [82].

Vector addition can be used to calculate the maximum theoretical cancellation that can be achieved by the amplitude and phase mismatch, though the control networks behind the adjustment units are generally quite complex, and subject to much secrecy within companies developing them for microwave amplifier systems.

For the system investigated in this work, where the feedforward correction takes place at a separate location to the error signal formation, some sort of feedback to the error signal formation will be required. In any system there will, of course, be a down-link and up-link path. Within this, gain/phase adjustment information could be included at the transmitter, to be processed by the error signal unit at the transmitter/receiver unit at the other end of the optical link. As any amplitude and phase degradation of the laser diodes will be slow, this feedback information does not have to be at high rates, and can hence be relatively simple in implementation, as long as it does not interfere with the transmitted RF signal.

Clearly, there is much that can be done to make this system a commercial reality, though this work has shown that the basic premise of feedforward non-linear correction is very viable. Ultimately, IMD compensation may probably be best achieved by a quasi-feedforward system, once matched laser diodes can be fabricated, with those characteristics degrading in a correspondingly identical manner. However, until that time, it is the conclusion of this work that a feedforward system, combined with a pre-distorted correction path laser, appears the most suitable solution.

References

1. *History of the Telephone*. 1998, Encyclopaedia Britannica online.
2. Gallagher, M. and W. Webb, *UMTS - the next generation of mobile radio.* Iee Review, 1999. **45**(2): p. 59-63.
3. Prasad, N.R., *GSM Evolution towards Third Generation UMTS/IMT 2000.* IEEE International Conterence on Personal Wireless Communications, 1999: p. 50-54.
4. Zeng, M.A., A.; Bhargava, Vijay K., *Recent advances in cellular wireless communications.* IEEE Communications Magazine, 1999. **37**(9): p. 128-138.
5. Pandya, R., *Mobile and Personal Communication Services and Systems*. Digital & Mobile Communicatio, ed. J.B. Anderson. 2000, New York: IEEE Press.
6. Mohr, W. and W. Konhauser, *Access network evolution beyond third generation mobile communications.* Ieee Communications Magazine, 2000. **38**(12): p. 122-133.
7. Costello, T. *Mobile - GSM/GPRS*. in *Telecoms Networks: The Next Generation*. 1999. University of Warwick, UK: IEE.
8. Breyer, S., et al., *A global view of the UMTS concept.* Alcatel Telecommunications Review, 1999(3): p. 219-227.
9. Matheson, G. *Mobile - UMTS*. in *Telecoms Networks: The Next Generation*. 1999. University of Warwick, UK: IEE.
10. BBC, *BBC Digital Radio*. 2003, BBC - Online: London. p. Interactive website.
11. Keller, H. and H. Salzwedel. *Link strategy for the mobile satellite system Iridium*. in *Proceedings of Vehicular Technology Conference - VTC, 28 April-1 May 1996*. 1996. Atlanta, GA, USA: IEEE.
12. Diaz, P., *The RAINBOW Concept for the UMTS Access Network.* IEEE International Symposium on Personal Indoor and Mobile Radio Communications, 1998. **1**: p. 429-433.
13. Toskala, A., H. Holma, and P. Muszynski. *ETSI WCDMA for UMTS*. in *Proceedings of ISSSTA'98 International Symposium on Spread Spectrum Techniques and Applications, 2-4 Sept. 1998*. 1998. Sun City, South Africa: IEEE.
14. Diaz, P., et al., *The RAINBOW Concept for the UMTS Access Network.* IEEE International Sympsosium on Personal Indoor and Mobile Radio Communications, 1998. **1**: p. 429-433.
15. Berruto, E., et al., *Research activities on UMTS radio interface, network architectures, and planning.* Ieee Communications Magazine, 1998. **36**(2): p. 82-95.
16. Mohr, W. *ACTS FRAMES project towards IMT-2000/UMTS*. in *1999 IEEE International Conference on Personal Wireless Communications, 17-19 Feb. 1999*. 1999. Jaipur, India: IEEE.
17. Nikula, E., et al., *FRAMES multiple access for UMTS and IMT-2000.* Ieee Personal Communications, 1998. **5**(2): p. 16-24.
18. Ojanpera, T., A. Klein, and P.-O. Anderson. *FRAMES multiple access for UTMS*. in *IEE Colloquium on CDMA Techniques and Applications for Third Generation Mobile Systems (Ref. No.1997/129), 19 May 1997*. 1997. London, UK: IEE.

19. Viterbi, A.J., *CDMA: Principles of Spread Spectrum Communication.* 1995, Wokingham, Reading, Mass.: Addison-Wesley.
20. Haberland, B., *UMTS radio demonstrator.* Alcatel Telecommunications Review, 1998(4): p. 292-297.
21. Chaudhury, P., W. Mohr, and S. Onoe, *The 3GPP proposal for IMT-2000.* Ieee Communications Magazine, 1999. **37**(12): p. 72-81.
22. Samukic, A., *UMTS universal mobile telecommunications system: Development of standards for the third generation.* Ieee Transactions on Vehicular Technology. 1998. **47**(4): p. 1099-1104.
23. Ojanpera, T. and R. Prasad, *An overview of third-generation wireless personal communications: A European perspective.* Ieee Personal Communications, 1998. **5**(6): p. 59-65.
24. Clapton, A.J., et al., *UMTS - the mobile part of broadband communications for the next century.* Bt Technology Journal, 1998. **16**(2): p. 120-131.
25. Toskala, A., et al., *FRAMES FMA2 Wideband-CDMA for UMTS.* European Transactions on Telecommunications, 1998. **9**(4): p. 325-336.
26. Dahlman, E., et al., *WCDMA - The radio interface for future mobile multimedia communications.* Ieee Transactions on Vehicular Technology, 1998. **47**(4): p. 1105-1118.
27. Dahlman, E., et al., *UMTS/IMT-2000 based on wideband CDMA.* Ieee Communications Magazine, 1998. **36**(9): p. 70-80.
28. Ojanpera, T. and R. Prasad, *An overview of air interface multiple access for IMT-2000/UMTS.* Ieee Communications Magazine, 1998. **36**(9): p. 82 - 95.
29. Yacoub, M.D. and R. Steele, *Chapters 21 and 22 from: The Mobile Communications Handbook, 2nd Ed.*, ed. J.D. Gibson. 2000, Boca Baton, Fl: CRC Press.
30. Lin, H.T. and Y.H. Kao, *Nonlinear distortions and compensations of DFB laser diode in AM-VSB lightwave CATV applications.* Journal of Lightwave Technology, 1996. **14**(11): p. 2567-2574.
31. Iwai, T., K. Sato, and K. Suto, *Signal Distortion and Noise in Am-Scm Transmission-Systems Employing the Feedforward Linearized Mqw-Ea External Modulator.* Journal of Lightwave Technology, 1995. **13**(8): p. 1606-1612.
32. Hung, W., et al., *Optical sampled subcarrier multiplexing scheme for nonlinear distortion reduction in lightwave CATV networks (vol 38, pg 1702, 2002).* Electronics Letters, 2003. **39**(25): p. 1873-1873.
33. Nakamura, N., et al., *Amplitude probability distribution of intermodulation distortion in multichannel digital optical cable transmission.* IEICE Transactions on Electronics

Recent Progress in Optoelectronics and Communications, 12-16 July 1998, 1999. **E82-C**(8): p. 1420-7.

34. Senior, J.M., *Optical Fibre Communications, principles and practice*. 1 ed. Prentice-Hall International Series in Optoelectronics, ed. P.J. Dean. 1985, London: Prentice/Hall International. 558.
35. Careless, J., *Signals from ground zero*. 2001, MRT - Mobile Radio Technology: New York. p. Website.

36. Junnarkar, S., *Network lessons from Ground Zero*. 2002, ZDNet: New York. p. Website.
37. Gu, X. and L.C. Blank, *10Gbit/s unrepeatered three-level optical transmission over 100km of standard fibre.* Electronics Letters, 1993. **29**(25): p. 2209-2211.
38. Rosher, P.A. and S.C. Fenning, *Multichannel video transmission over 100km of step index singlemode fibre using a directly modulated 1537nm distributed feedback laser.* Electronics Letters, 1990. **26**(8): p. 534-536.
39. Favre, F., et al., *320Gbit/s soliton WDM transmission over 1300km with 100km dispersion-compensated spans of standard fibre.* Electronics Letters, 1997. **33**(25): p. 2135-2136.
40. *Radio Optical Broadband Integrated Network, 'ROBIN', Final Report.* 1999, University of Bradford, Nortel, Orange: Bradford/Harlow, UK.
41. *A System Designer's Guide to RF and Microwave Fiber Optics*, Ortel Corporation (part of Agere Systems): Allentown, PA. p. 4-16.
42. Muirhead, R.A., *Theory and Design of Radio Over Fibre Links*, in *Department of Electronic Imaging and Media Communications*. 2000, University of Bradford: Bradford.
43. *Optical Fiber Telecommunications IIIA*, ed. I.P. Kaminow and T.L. Koch. 1997, San Diego/London: Academic Press.
44. Haykin, S., *Communication Systems, 4th Ed.* 2001, New York/Chichester: Wiley.
45. *Optical Fiber Telecommunications IIIB*, ed. I.P. Kaminow and T.L. Koch. 1997, San Diego/London: Academic Press.
46. Hassin, D. and R. Vahldieck, *Feedforward Linearization of Analog Modulated Laser-Diodes - Theoretical-Analysis and Experimental-Verification.* Ieee Transactions on Microwave Theory and Techniques, 1993. **41**(12): p. 2376-2382.
47. Fock, L.S., A. Kwan, and R.S. Tucker, *Reduction of Semiconductor-Laser Intensity Noise by Feedforward Compensation - Experiment and Theory.* Journal of Lightwave Technology, 1992. **10**(12): p. 1919-1925.
48. Moyer, M., *Quasi-Feedforward Compensation of Laser Diode Distortion*. 1998, McMaster University: Hamilton, Ontario.
49. Grotzinger, T.L., *The effects of atmospheric conditions on the performance of free-space infrared communications.* SPIE Free-Space Laser Communications Technologies III, 1991. **1417**.
50. Kim, I.I., K. McArthur, and E. Korevaar, *Comparison of laser beam propagation at 785nm and 1550nm in fog and haze for optical wireless communications.* SPIE Optical Wireless Communications III, 2001. **4214**.
51. *Blue Sky Research Announces Laser Diode Sources for Free Space Optical Communications*. 2001, Blue Sky Research. p. Web page.
52. Liang, K., *Nonlinear Characterisation of Quantum Well Lasers and Linearisation by Predistortion for 1.8GHz Narrow Band Optical Transmitters*, in *Department of Electronic Imaging and Media Communications*. 2000, University of Bradford: Bradford.
53. Volterra, V., *Theory of functionals (and of integral and integro-differential equations)*. 1931, London/Glasgow: Blackie.
54. Pan, Q. and R.J. Green, *Pre-Clipping AM/QAM Hybrid Lightwave Systems with Bandstop Filtering.* IEEE Photonics Technology Letters.

55. Pan, Q. and R.J. Green, *Performance analysis of preclipping AM/QAM hybrid lightwave systems.* Journal of Lightwave Technology, 1997. **15**(1): p. 1-5.
56. Cripps, S.C., *RF Power Amplifiers for Wireless Communications*. Artech House Microwave Library. 1999, Boston/London: Artech House.
57. Van De Grijp, A. and e. al.. *Novel electro-optical feedback for noise and distortion reduction in high-quality analogue optical transmission video signals.* Electronics Letters, 1981. **17**(11): p. 361-362.
58. Straus, J. and e. al.. *Phase-shift modulation technique for the linearisation of analogue optical transmitters.* IEE Electronic Letters, 1977. **13**(5): p. 149-151.
59. Fock, L.S. and R.S. Tucker, *Simultaneous Reduction of Intensity Noise and Distortion in Semiconductor-Lasers by Feedforward Compensation.* Electronics Letters, 1991. **27**(14): p. 1297-1299.
60. Straus, J. and O.I. Szentes, *Linearisation of optical transmitters by a quasifeedforward compensation technique.* Electronics Letters, 1977. **13**(6): p. 158-9.
61. Stroud, K.A., *Engineering Mathematics, 5th Ed.* 2001: Palgrave.
62. Highman, D.J. and N.J. Highman, *Matlab Guide*. 2000, Philadelphia, PA: Siam.
63. Yariv, A., *Optical Electronics*. Forth ed. 1991: Saunders College Publishing.
64. Betti, S., E. Bravi, and M. Giaconi, *Analysis of the performance of subcarrier multiplexed (SCM) optical systems.* Computer Networks-the International Journal of Computer and Telecommunications Networking, 2000. **32**(5): p. 563-569.
65. Helms, J., *Intermodulation Distortions of Broad-Band Modulated Laser- Diodes.* Journal of Lightwave Technology, 1992. **10**(12): p. 1901-1906.
66. Wiener, N., *Response of a nonlinear device to noise.* M.I.T Radiation Lab., Rep.. 1942. **16S**.
67. Salgado, H.M. and J.J. O'Reilly, *Volterra series analysis of distortion in semiconducter laser diodes.* IEE Proceedings (Part J) Optoelectronics, 1991. **138**(6): p. 379-382.
68. Salgado, H.M. and J.J. O'Reilly, *Experimental validation of Volterra series nonlinear modelling for microwave subcarrier optical systems.* IEE Proceedings Optoelectronics, 1996. **143**(4): p. 209-213.
69. Biswas, T.K. and W.F. McGee, *Volterra Series Analysis of Semiconductor-Laser Diode.* Ieee Photonics Technology Letters, 1991. **3**(8): p. 706-708.
70. Hunziker, S., *Volterra analysis of second- and third-order intermodulation of InGaAsP/InP laser diodes: theory and experiment.* Optical Engineering, 1995. **34**(7): p. 2037-43.
71. Schetzen, M., *The Volterra and Wiener theories of nonlinear systems*. 1980, New York/Chichester: Wiley.
72. Tucker, R.S. and D.J. Pope, *Circuit Modeling of the Effect of Diffusion on Damping in a Narrow-Stripe Semiconductor-Laser.* Ieee Journal of Quantum Electronics, 1983. **19**(7): p. 1179-1183.
73. Bedrosian, E. and S.O. Rice, *The output properties of Volterra systems (nonlinear systems with memory) driven by harmonic and gaussian inputs.* Proceedings IEEE, 1971. **59**(12): p. 1688-1707.

74. Sweet, C.S. and R.J. Green. *Optical Free Space Feedforward Non-linear Correction System for Mobile RF*. in *Wireless Design Conference*. 2002. London, UK: IEEE and Microwave Engineering.
75. Nagarajan, R., et al., *Effects of Carrier Transport on High-Speed Quantum-Well Lasers.* Applied Physics Letters, 1991. **59**(15): p. 1835-1837.
76. Nagarajan, R., et al., *High-Speed Quantum-Well Lasers and Carrier Transport Effects.* Ieee Journal of Quantum Electronics, 1992. **28**(10): p. 1990-2008.
77. Tsai, C.Y., et al., *Nonlinear Gain Coefficients in Semiconductor Quantum-Well Lasers - Effects of Carrier Diffusion, Capture, and Escape.* Ieee Journal of Selected Topics in Quantum Electronics, 1995. **1**(2): p. 316-330.
78. Vassilovski, D., et al., *Unambiguous Determination of Quantum Capture, Carrier Diffusion, and Intrinsic Effects in Quantum-Well Laser Dynamics Using Wavelength-Selective Optical Modulation.* Ieee Photonics Technology Letters, 1995. **7**(7): p. 706-708.
79. Fock, L.S. and R.S. Tucker, *Reduction of Distortion in Analog Modulated Semiconductor- Lasers by Feedforward Compensation.* Electronics Letters, 1991. **27**(8): p. 669-671.
80. Ramirez-Inguez, R. and R.J. Green, *Totally internally reflecting optical antennas for wireless IR communication.* IEEE Wireless Design Conference, 2002: p. 129-132.
81. Ramirez-Inguez, R., *High Gain Front-Ends for Optical Wireless Receivers*, in *School of Engineering*. 2002, Warwick: Coventry.
82. Pothecary, N., *Foodforward Linear Power Amplifiers*. Artech House Microwave Library. 1999, Boston/London: Artech House.

Appendix A: Perturbation analysis

The details of the small signal derivation by Helms for LD intermodulation distortion with definitions of the relavent functions is shown in this appendix.

Using the rate equations of (3.5) and (3.6):

$$\frac{dN}{dt} = \frac{I}{qV} - \frac{N}{\tau_n} - G(N,Q)Q \qquad \textbf{(3.5), A1}$$

$$\frac{dQ}{dt} = \Gamma G(N,Q)Q - \frac{Q}{\tau_p} + \beta \qquad \textbf{(3.6), A2}$$

where Q is the photon density, N is the electron density, I is the current injected into the active region, V is the volume of the active region, q is the charge of an electron, τ_n is the recombination lifetime of the carriers (or lifetime of the excited state), τ_p is the photon lifetime, Γ is the optical confinement factor, β is the spontaneous emission rate and $G(N,Q) = g(N-N_o)(1-\varepsilon Q)$, is the material gain, where ε is the gain suppression coefficient, N_o is the transparency carrier density and g is the differantial gain.

The expressions for I(t), N(t) and Q(t) for an LD with a sinusoidal current input are thus as follows:

$$I(t) = I_{bias} + Re(\Delta I e^{j\omega t}) \qquad \textbf{(3.10), A3}$$

$$N(t) = N_{th} + Re(\Delta N e^{j\omega t}) \qquad \textbf{(3.11), A4}$$

$$Q(t) = Q_m + Re(\Delta Q e^{j\omega t}) \qquad \textbf{(3.12), A5}$$

These expressions can then be fed back into the rate equations of A1 and A2. By neglecting any terms higher than the 3rd order and the noise sources a set of linearised equations in the frequency domain can be found:

$$\left(j\omega + \frac{\gamma}{\tau_p}\right)\Delta Q(j\omega) - gQ_m \Delta N(j\omega) = D_Q(j\omega) \qquad \textbf{A6}$$

$$\frac{1}{V\tau_p}\Delta Q(j\omega) + \left(j\omega + \frac{\gamma}{\tau_p}\right)\Delta N(j\omega) = D_N(j\omega) \qquad \textbf{A7}$$

where $D_Q(j\omega)$ and $D_N(j\omega)$ are the driving terms, and the damping rate, γ and the effective carrier lifetime, τ_r are defined thus:

$$\gamma = (\tau_p + \varepsilon/g)\omega_r \qquad \textbf{A8}$$

$$1/\tau_r = 1/\tau_n + \omega_r^2 \tau_p \qquad \textbf{A9}$$

where ω_r is the relaxation oscillation frequency that, using the steady state solution to Equ. A2, can be manipulated into the following expression:

$$\omega_r = \sqrt{\frac{1}{\tau_p \tau_n}\left(\frac{I}{I_{th}} - 1\right)} \qquad \textbf{A9}$$

The first order solutions to Equ. A6 and A7 (denoted subscript 1) are acquired using the first order driving terms:

$$D_{Q,1}(j\omega) = 0 \quad \text{and} \quad D_{N,1}(j\omega) = \Delta I(j\omega)/qV \qquad \textbf{A10}$$

with the higher order solutions iteratively found from Equ. A6 and A7 using the lower order solutions to give an overall solution, with the k^{th} order driving terms, as:

$$D_{Q,k}(j\omega_k) = \frac{g}{2}\sum \Delta Q_i(j\omega_i)\Delta N_j(j\omega_j) \qquad \textbf{A11}$$

$$D_{N,k}(j\omega_k) = -D_{Q,k}(j\omega_k) \qquad \textbf{A12}$$

where the summation is over all terms with:

$$k = i + j \quad \text{and} \quad \omega_k = \pm\omega_i \pm \omega_j \qquad \textbf{A13}$$

As stated in chapter 3, the LD's normalised transfer function is:

$$H(j\omega) = \frac{1}{(j\omega/\omega_r)^2 + j\omega/j\omega_d + 1} \qquad \textbf{(3.13), A14}$$

where ω_r, the relaxation oscillation frequency, is defined as in equ. A9, and ω_d is the damping frequency, given by:

$$\omega_d = \frac{\omega_r^2}{\frac{\gamma}{\tau_p} + \frac{1}{\tau_r}} \qquad \textbf{A15}$$

The optical modulation index, m_{opt}, is usually assumed to be close to the value of the current modulation index, albeit marginally larger. For the perturbation analysis m_{opt} is defined as follows:

$$m_{opt}(j\omega) = H(j\omega)\frac{\Delta I(j\omega)\tau_p}{qQ_m} \qquad \textbf{A16}$$

However, a more user friendly definition can be found in equation (3.28) in section 3.3 as part of the Volterra analysis.

The two and three-tone intermodulation products that make up equations (3.15) and (3.16) are found from the following ratios:

$$ID2 = \left|\frac{\Delta Q_2(j\omega_{ID2})}{\Delta Q(j\omega_1)}\right| \qquad \textbf{A17}$$

$$ID3 = \left|\frac{\Delta Q_3(j\omega_{ID3})}{\Delta Q(j\omega_1)}\right| \qquad \textbf{A18}$$

Where $\omega_{ID2} = \omega_1 \pm \omega_2$ and $\omega_{ID3} = \omega_1 \pm \omega_2 \pm \omega_3$.

Equation (3.16) of section 3.2 of chapter 3 gives an expression for the third order intermodulation products, IMD_3. From this the expression for $C(j\omega')$ is as such:

$$C(j\omega') = \frac{1}{\omega_r}\left[H(j\omega_{1,2})A(j\omega_{1,2})\left(-j\omega_3 + \frac{\gamma}{\tau_p}\right) + H(j\omega_{1,3})A(j\omega_{1,3})\left(j\omega_2 + \frac{\gamma}{\tau_p}\right)\right.$$

$$\left. + H(j\omega_{2,3})A(j\omega_{2,3})\left(j\omega_1 + \frac{\gamma}{\tau_p}\right)\right] - \omega_r\tau_p\left[H(j\omega_{1,2})B(j\omega_{1,2})\right.$$

$$\left. + H(j\omega_{1,3})B(j\omega_{1,3}) + H(j\omega_{2,3})B(j\omega_{2,3})\right| \qquad \textbf{A19}$$

Where, for a two tone input case, $\omega_{1,2} = \omega_1 + \omega_2 \Rightarrow 2\omega_1$; $\omega_{1,3} = \omega_1 - \omega_3 \Rightarrow \omega_1 - \omega_2$; $\omega_{2,3} = \omega_2 - \omega_3 \Rightarrow \omega_1 - \omega_2$.

The optical mudulation indexes, m_{opt}, for the two signals are assumed to be equal.

The expressions for $A(j\omega)$ and $B(j\omega)$ are as shown below:

$$A(j\omega) = \frac{2\gamma}{\tau_p\tau_n\omega_r^2} - \left(\frac{\omega}{\omega_r}\right)^2 + \frac{j\omega}{\omega_r^2}\left(\frac{1}{\tau_n} + \frac{2\gamma}{\tau_p}\right) \qquad \textbf{A20}$$

$$B(j\omega) = \frac{2\gamma}{\tau_p\omega_r^2} - \left(\frac{\omega}{\omega_r}\right)^2 + \frac{j\omega}{\tau_p\omega_r^2} \qquad \textbf{A21}$$

For the full and complete explanation see Helms [35].

Appendix B: Volterra Analysis

Appendix B1: Standard Analysis.

As stated in the main text, this appendix serves as a more comprehensive summary of the Volterra derivation, rather than a full in depth analysis, which can be found as referenced in section 3.3.

Again, starting with the LD rate equations of (3.5) and (3.6) as shown below:

$$\frac{dN}{dt} = \frac{I}{qV} - \frac{N}{\tau_n} - G(N,Q)Q \quad \text{and} \quad \frac{dQ}{dt} = \Gamma G(N,Q)Q - \frac{Q}{\tau_p} + \beta \qquad \textbf{B1.1}$$

where as before, Q is the photon density, N is the electron density, I is the current injected into the active region, V is the volume of the active region, q is the charge of an electron, τ_n is the recombination lifetime of the carriers (or lifetime of the excited state), τ_p is the photon lifetime, Γ is the optical confinement factor, β is the spontaneous emission rate. However, this time expanding upon the term for material gain gives $G(N,Q) = g'(N-N_0)$ where g' is the differential gain and N_0 is the transparency carrier density. The differential gain can be further expanded to, $g' = g(1-\varepsilon Q)$, giving the overall expression as:

$$G(N,Q) = g(1-\varepsilon Q)(N-N_0) \qquad \textbf{(3.16), B1.2}$$

Where g is the optical gain factor, which is determined material factors (and could be dependent upon the density of carriers). The term ε is the gain suppression parameter and is a phenomenological constant, which identifies the characteristics of the gain compression of the active region.

Again, with the laser biasing point well above the threshold condition, the spontaneous process can be effectively ignored with respect to the stimulation

process, i.e. giving $\beta = 0$, the two rate equations of B1.1 can be combined to form a third equation that expresses the injection current I in terms of the photon density as can be seen below:

$$I - I_{th} = \frac{V'}{\Gamma}\left(\frac{dQ}{dt} + \frac{Q}{\tau_p}\right) + \frac{V'}{\Gamma g}\frac{d}{dt}\left[\frac{\frac{dQ}{dt} + \frac{Q}{\tau_p}}{(1-\varepsilon Q)Q}\right] \qquad \textbf{(3.17), B1.3}$$

where $V' = qV$, $I_{th} = qVN_0/\tau_n$ and, due to the large value of τ_n, the $1/\tau_n$ term has been set to zero for simplicity.

Now both the injection current and photon density are separated into ac and dc parts:

$$I(t) = I_0 + i(t) \quad \text{and} \quad Q(t) = Q_0 + q(t) \qquad \textbf{(3.18), B1.4}$$

where I_0 is the laser biasing current, i(t) is the modulation current, Q_0 is the steady state photon density associated with I_0, and q(t) is the time varying photon density associated with i(t). Next, the denominator of the last term in B1.3 is substituted by its Taylor's series expansion around the steady state, as shown below:

$$f(Q) = \frac{1}{(1-\varepsilon Q)Q}$$

$$\cong f(Q_0) + f'(Q_0)q(t) + f''(Q_0)q^2(t)/2! + f'''(Q_0)q^3(t)/3! \qquad \textbf{B1.5}$$

where

$$f(Q_0) = \frac{1}{1-\varepsilon Q_0)Q_0} \qquad \textbf{B1.6}$$

$$f'(Q_0) = (2\varepsilon Q_0 - 1)f^2(Q_0) \qquad \textbf{B1.7}$$

$$f''(Q_0) = 2(3\varepsilon^2 Q_0^2 - 3\varepsilon Q_0 + 1)f^3(Q_0) \qquad \textbf{B1.8}$$

$$f'''(Q_0) = 6(4\varepsilon^3 Q_0^3 - 6\varepsilon^2 Q_0^2 - 4\varepsilon Q_0 - 1)f^4(Q_0) \qquad \textbf{B1.9}$$

Substituting equations B1.4 to B1.9 into B1.3 then gives:

$$i(t) = A + Dq(t) + Eq'(t) + Fq''(t) + Mq(t)q'(t) + Nq(t)q''(t) + Nq'^{2}(t) + Sq^{2}(t)q'(t) + 2Gq(t)q'^{2}(t) + Gq^{2}(t)q''(t) + \ldots. \quad \textbf{(3.19), B1.10}$$

where A to G are constants and are expressed in terms of the LD parameters, as shown below in equations B1.11 to B1.18:

$$A = I_{th} - I_0 + \frac{V'}{\Gamma\tau_p}Q_0 = 0, \text{ and } Q_0 = \frac{\Gamma\tau_p}{V'}(I_0 - I_{th}) \quad \textbf{B1.11}$$

$$D = \frac{V'}{\Gamma\tau_p} \quad \textbf{B1.12}$$

$$E = \frac{V'}{\Gamma}[1 + \frac{f(Q_0) + Q_0 f'(Q_0)}{g\tau_p}] \quad \textbf{B1.13}$$

$$F = \frac{V'}{\Gamma g}f(Q_0) \quad \textbf{B1.14}$$

$$M = \frac{V'}{\Gamma g\tau_p}[2f'(Q_0) + Q_0 f''(Q_0)] \quad \textbf{B1.15}$$

$$N = \frac{V'}{\Gamma g}f'(Q_0) \quad \textbf{B1.16}$$

$$S = \frac{V'}{2\Gamma g\tau_p}[3f''(Q_0) + Q_0 f'''(Q_0)] \quad \textbf{B1.17}$$

$$N = \frac{V'}{2\Gamma g}f''(Q_0) \quad \textbf{B1.18}$$

The Volterra series can now be used to relate i(t) and q(t) in the non-linear distortion modelling, as follows:

$$i(t) = \sum_{n=1}^{\infty}\frac{1}{n!}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}g_n(u_1\ldots u_n)\cdot\prod_{k=1}^{n}q(t-u_k)du_1\ldots du_n \quad \textbf{(3.20), B1.19}$$

in which, $g_n(u_1\ldots u_n)$ is called the *n*th order Volterra kernels of the system. Equation B1.19 gives the true input current in terms of the true output and is called the inverse

laser system. To acquire the forward Volterra transfer functions, $H_n(\omega_1,\ldots\omega_n)$, which treat i(t) as the input of the system, the inverse Volterra transfer functions, $G_n(\omega_1,\ldots\omega_n)$, for the system which treat q(t) as their inputs, have to be derived first.

Using the harmonic input method, the first three G_n are calculated from equation B1.10. If $q(t) = e^{j\omega t}$, then equation B1.19 becomes

$$i(t) = \int_{-\infty}^{\infty} g_1(\tau)\cdot e^{j\omega(t-\tau)}d\tau = G_1(\omega)\cdot e^{j\omega t} \qquad \textbf{B1.20}$$

Substituting q(t) and equation B1.20 into equation B1.10, and taking the coefficients of $e^{j\omega t}$ results in the first order Volterra transfer function, shown below:

$$G_1(\omega) = D - F\omega^2 + j\omega E \qquad \textbf{(3.21), B1.21}$$

The second order Volterra transfer function is derived by substituting $q(t) = e^{j\omega 1t} + e^{j\omega 2t}$ into B1.10 and taking the coefficients of $e^{j(\omega 1 + \omega 2)t}$, giving:

$$G_2(\omega_1,\omega_2) = j(\omega_1 + \omega_2)M - (\omega_1 + \omega_2)^2 N \qquad \textbf{(3.22), B1.22}$$

Similarly, the third order term is acquired by substituting $q(t) = e^{j\omega 1t} + e^{j\omega 2t} + e^{j\omega 3t}$ and taking the coefficients of the term $e^{j(\omega 1 + \omega 2 + \omega 3)t}$, giving

$$G_3(\omega_1,\omega_2,\omega_3) = j2S(\omega_1 + \omega_2 + \omega_3) - 2G(\omega_1 + \omega_2 + \omega_3)^2 \qquad \textbf{(3.23), B1.23}$$

From the above, using the *n*th order inverse and the harmonic balance method, the forward Volterra transfer functions required for actual system modelling, $H_n(\omega_1,\ldots\omega_n)$, can be calculated from G_n. The first order functions have a simple relationship, as in linear systems:

$$H_1(\omega) = \frac{1}{G_1(\omega)} = \frac{1}{D + j\omega E - F\omega^2} \qquad \textbf{(3.24), B1.24}$$

The 2nd order function, $H_2(\omega_1,\omega_2)$ is acquired by applying an input equal to $x(t) = e^{j\omega 1t} + e^{j\omega 2t}$ to the system and cancelling the $e^{j(\omega 1 + \omega 2)t}$ terms corresponding to the operation of the 2nd order system. The output of the first 2nd order system, y(t) would be:

$$y(t) = G_1(\omega_1)e^{j\omega 1t} + G_2(\omega_2)e^{j\omega 2t} + 2G_2(\omega_1,\omega_2)e^{j(\omega 1 + \omega 2)t} \qquad \textbf{B1.25}$$

This is considered as the input to another 2nd order system, as illustrated by Fig. B1.1, giving a final output of:

$$\begin{aligned} z(t) = {} & H_1(\omega_1)G_1(\omega_1)e^{j\omega 1t} + H_1(\omega_2)G_1(\omega_2)e^{j\omega 2t} \\ & + H_1(\omega_1 + \omega_2)2G_1(\omega_1,\omega_2)e^{j(\omega 1 + \omega 2)t} \\ & + 2H_2(\omega_1,\omega_2)G_1(\omega_1)G_1(\omega_2)e^{j(\omega 1 + \omega 2)t} \\ & + 2H_2(\omega_1,\omega_1 + \omega_2)2G_1(\omega_1)G_2(\omega_1,\omega_2)e^{j(2\omega 1 + \omega 2)t} \\ & + 2H_2(\omega_2,\omega_1 + \omega_2)2G_1(\omega_2)G_2(\omega_1,\omega_2)e^{j(\omega 1 + 2\omega 2)t} \\ & +\ldots, \end{aligned} \qquad \textbf{B1.26}$$

The coefficients of $e^{j(\omega 1 + \omega 2)t}$ should be zero, for the overall system to act as a linear system, therefore:

$$H_1(\omega_1 + \omega_2)G_2(\omega,\omega_2) + H_2(\omega_1,\omega_2)G_1(\omega_1)G_1(\omega_2) = 0 \qquad \textbf{B1.27}$$

Combining equations B1.27 and B1.24 gives the 2nd order transfer function, as seen below:

$$H_2(\omega_1,\omega_2) = -\frac{G_2(\omega_1,\omega_2)}{G_1(\omega_1)G_1(\omega_2)G_1(\omega_1+\omega_2)} \qquad \textbf{B1.28}$$

The 3rd order transfer function is then acquired in a similar way, with a three tone input of, $x(t) = e^{j\omega 1t} + e^{j21t} + e^{j\omega 3t}$ resulting in the output of the first system, y(t), that is then the input into the cascade system. The coefficients of the $e^{j(\omega 1 + \omega 2 + \omega 3)t}$ term in the output, z(t), should again be zero, resulting in the following term for the 3rd order transfer function:

$$H_3(\omega_1,\omega_2,\omega_3) = \frac{-1}{G_1(\omega_1)G_1(\omega_2)G_1(\omega_3)G_1(\omega_1+\omega_2+\omega_3)} \cdot \Big[G_3(\omega_1,\omega_2,\omega_3) - \frac{G_2(\omega_1,\omega_2+\omega_3)G_2(\omega_2,\omega_3)}{G_1(\omega_2+\omega_3)} - \frac{G_2(\omega_2,\omega_1+\omega_3)G_2(\omega_1,\omega_3)}{G_1(\omega_1+\omega_3)} - \frac{G_2(\omega_3,\omega_1+\omega_2)G_2(\omega_1,\omega_2)}{G_1(\omega_1+\omega_2)}\Big] \qquad \textbf{B1.29}$$

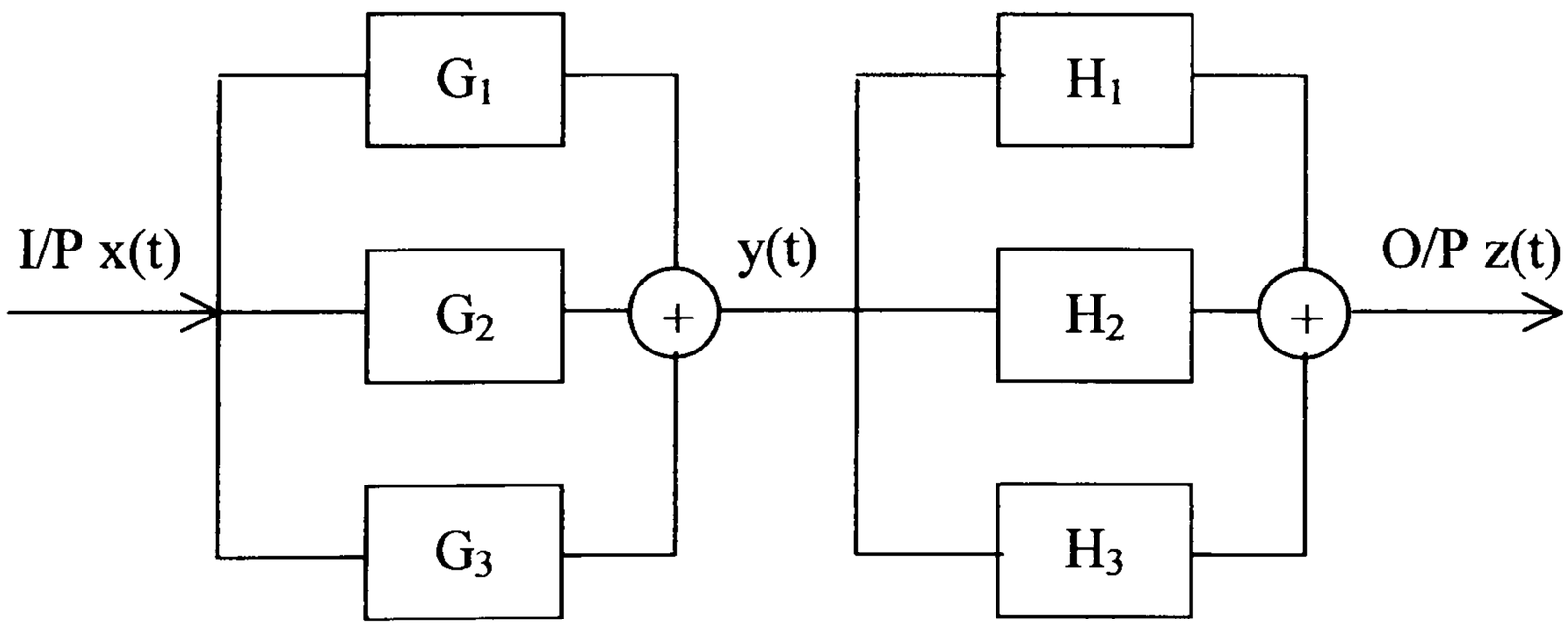


Fig. B1.1 *Block diagram of a cascade of two systems inverse one another modelled in Volterra series*

The Volterra series expresses the output, z(t) of a non-linear system in 'powers' of the input, x(t). Typically, the series can be written as:

$$z(t) = \sum_{n=1}^{\infty} \int_{-\infty}^{\infty} d\tau_1 .... \int_{-\infty}^{\infty} d\tau_n h_n(\tau_1,...,\tau_n) \prod_{r=1}^{n} x(t-\tau_r) \qquad \textbf{B1.30}$$

where the kernels $h_n(\tau_1,\ldots, \tau_n)$ describe the system, and the 1st order kernel, $h_1(\tau_1)$ is the impulse response of a linear network. Hence, the higher order kernels can be viewed as higher order impulse responses, serving to characterise the various orders of non-linearity. Also, the Fourier transform $H_n(f_1,\ldots, f_n)$ of the nth-order Volterra kernel is as an nth-order Volterra transform function. When the $H_n(f_1,\ldots, f_n)$ are known, relevant terms of the output can be acquired by substituting the $H_n(f_1,\ldots, f_n)$ in general formulas derived from the Volterra series representation. Thus, for a two sine wave input, $x(t) = P\cos\omega_1 t + Q\cos\omega_2 t$, where $\omega_{1,2} = 2\pi f_{1,2}$, the leading terms in the dc and lower order components of the output are as follows [43]:

$$\left[\frac{P^2}{4}H_2(f_1,-f_1)+\frac{Q^2}{4}H_2(f_2,-f_2)\right]$$

$$e^{j\omega 1t}\left[\frac{P}{2}H_1(f_1)+\frac{Q^3}{16}H_3(f_1,f_1,-f_1)+\frac{PQ^2}{8}H_3(f_1,f_2,-f_2)\right],$$

$$e^{j2\omega 1t}\frac{P^2}{8}H_2(f_1,f_1),$$

$$e^{j(\omega 1+\omega 2)t}\frac{PQ}{4}H_2(f_1,f_2),$$

$$\text{and } e^{j(2\omega 1+\omega 2)t}\frac{P^2Q}{16}H_3(f_1,f_1,f_2). \qquad \textbf{B1.31}$$

Taking the components for the fundamental, the 2nd order intermodulation at frequency $(\omega_1 + \omega_2)$ and the 3rd order intermodulation at frequency $(2\omega_1 - \omega_2)$ from B1.31 above gives:

$$\text{Fund} = PH_1(\omega_1)/2 \qquad \textbf{(3.25), B1.32}$$

$$IMD_2 = P^2H_2(\omega_1,\omega_2)/4 \qquad \textbf{(3.26), B1.33}$$

$$IMD_3 = P^3H_3(\omega_1, \omega_1, -\omega_2)/16 \qquad \textbf{(3.27), B1.34}$$

These can now be implemented within Matlab and used within the main simulation program, as for the Taylor's series and Perturbation techniques.
A full explanation of this technique can be found in Liang [24].

Appendix B2: Advanced Analysis.

As stated in section 3.4, the analysis including the carrier transport effects follows much the same path as that for the standard analysis in section 3.3/Appendix B1 (AppB1). Also, as stated in AppB1, the following serves as a more complete summary, rather than a full analysis. Starting with the three rate equation model:

$$\frac{dQ}{dt} = Q\left(\Gamma G - \frac{1}{\tau_p}\right) \qquad \textbf{(3.29), B2.1}$$

$$\frac{dN_{qw}}{dt} = -\frac{N_{qw}}{\tau_n} - \frac{N_{qw}}{\tau_{esc}} + \frac{N_{sch}}{\tau_{cap}} \cdot \frac{V_{sch}}{V_{qw}} - QG \qquad \textbf{(3.30), B2.2}$$

$$\frac{dN_{sch}}{dt} = \frac{I}{eV_{sch}} - \frac{N_{sch}}{\tau_{cap}} + \frac{N_{qw}}{\tau_{esp}} \cdot \frac{V_{qw}}{V_{sch}} - \frac{N_{sch}}{\tau_n} \qquad \textbf{(3.31), B2.3}$$

where G is the material gain and can be written as $G = g(1 - \varepsilon Q)(N_{qw} - N_{qw0})$ and N_{qw0} is the carrier density for transparency, V_{qw} is the volume of the quantum wells, V_{sch} is the volume of the SCH and barrier regions, τ_{cap} is the carrier capture time into the quantum well, and τ_{esc} is the carrier escape time from the quantum well. Other symbols are defined as in AppB1.

Neglecting the recombination of the carrier in the SCH/barrier region, and setting $1/\tau_n$ to zero, as in AppB1, equations B2.1 to B2.3 can be reduced to;

$$I - I_{th} = \frac{V'}{\Gamma}\left(\frac{dQ}{dt} + \frac{Q}{\tau_p}\right) + \frac{V'(1+R_{rd})}{\Gamma g} \cdot \frac{d}{dt}\left[\frac{\frac{dQ}{dt} + \frac{Q}{\tau_p}}{(1-\varepsilon Q)Q}\right] \qquad \textbf{(3.32), B2.4}$$

$$+ \frac{\tau_{cap} V'}{\Gamma} \cdot \frac{d}{dt}\left(\frac{dQ}{dt} + \frac{Q}{\tau_p}\right) + \frac{\tau_{cap} V'}{\Gamma g} \frac{d^2}{dt^2}\left[\frac{\frac{dQ}{dt} + \frac{Q}{\tau_p}}{(1-\varepsilon Q)Q}\right]$$

where, $I_{th} = qV_{qw}N_{qw0}/\tau_n$ and $V' = qV_{qw}$. Also defined in B2.4 is the capture/escape time ratio, $R_{rd} = \tau_{cap}/\tau_{esc}$, which is also known as the transport factor.

Following the proceedure shown in AppB1, with $I = I_0 + i(t)$, and $Q = Q_0 + q(t)$, and substituting equations B1.4 to B1.9 into B2.4, it can be re-written in respect to the driving current as;

$$i(t) = A + B_1q(t) + B_2q'(t) + B_3q''(t) + B_4q'''(t) + C_1q(t)q'(t) + C_2q(t)q''(t) + C_2q(t)q''(t) + C_3q(t)q'''(t) + 3C_3q'(t)q''(t) + D_1q'(t)q^2(t) + D_2q'^2(t)q(t) + 0.5D_2q''(t)q^2(t) + D_3q'''(t)q^2(t) + 6D_3q'''(t)q''(t)q(t) + 2D_3q'^3(t) \quad \textbf{(3.33), B2.5}$$

where A, B_1 to B_4, C_1 to C_3 and D_1 to D_3 are defined by the laser intrinsic parameters and are defined as below:

$$A = I_{th} - I_0 + \frac{V'}{\Gamma\tau_p}Q_0 = 0 \quad \textbf{B2.6}$$

$$B_1 = \frac{V'}{\Gamma\tau_p} \quad \textbf{B2.7}$$

$$B_2 = \frac{V'}{\Gamma g\tau_p}(1+R_{rd})[Q_0f'(Q_0)+f(Q_0)] + \frac{V'\tau_{cap}}{\Gamma\tau_p} + \frac{V'}{\Gamma} \quad \textbf{B2.8}$$

$$B_3 = \frac{V'\tau_{cap}}{\Gamma g\tau_p}[Q_0f'(Q_0)+f(Q_0)] + \frac{V'}{\Gamma\tau_p}f(Q_0)(1+R_{rd}) + \frac{V'\tau_{cap}}{\Gamma} \quad \textbf{B2.9}$$

$$B_4 = \frac{V'\tau_{cap}}{\Gamma g}f(Q_0) \quad \textbf{B2.10}$$

$$C_1 = \frac{V'}{\Gamma g\tau_p}(1+R_{rd})[Q_0f''(Q_0)+2f'(Q_0)] \quad \textbf{B2.11}$$

$$C_2 = \frac{V'\tau_{cap}}{\Gamma g\tau_p}[Q_0f''(Q_0)+2f'(Q_0)] + \frac{V'}{\Gamma\tau_p}f'(Q_0)(1+R_{rd}) \quad \textbf{B2.12}$$

$$C_3 = \frac{V'\tau_{cap}}{\Gamma g}f'(Q_0) \quad \textbf{B2.13}$$

$$D_1 = \frac{V'}{2\Gamma g\tau_p}(1+R_{rd})[Q_0f'''(Q_0)+3f''(Q_0)] \quad \textbf{B2.14}$$

$$D_2 = \frac{V'\tau_{cap}}{\Gamma g\tau_p}[Q_0f'''(Q_0)+3f''(Q_0)] + \frac{V'}{\Gamma\tau_p}f''(Q_0)(1+R_{rd}) \quad \textbf{B2.15}$$

$$D_3 = \frac{V'\tau_{cap}}{2\Gamma g} f''(Q_0) \qquad \textbf{B2.16}$$

Again, following the a similar proceedure to that shown in AppB1 the first three-order inverse Volterra transfer functions can be derived, as shown below.

$$G_1(\omega) = B_1 - B_3\omega^2 + j(B_2\omega - B_4\omega^3) \qquad \textbf{(3.34), B2.17}$$

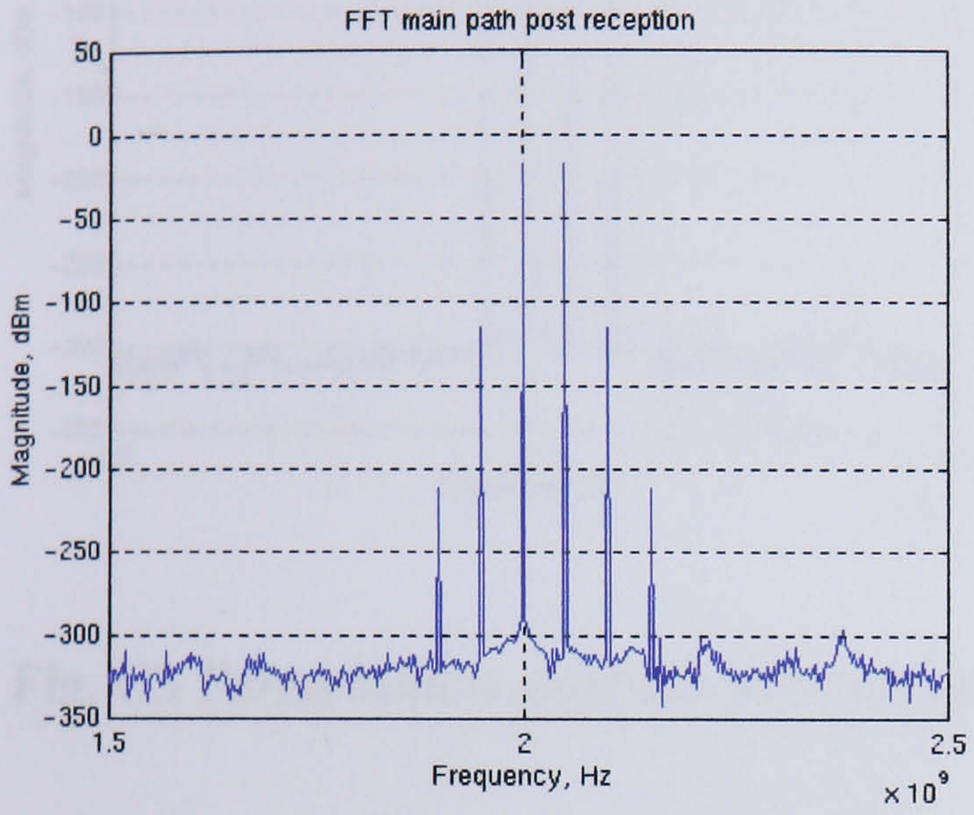
$$G_2(\omega_1,\omega_2) = -C_2(\omega_1 + \omega_2)^2 + j(\omega_1 + \omega_2)[C_1 - C_3(\omega_1 + \omega_2)^2] \qquad \textbf{(3.35), B2.18}$$

$$G_3(\omega_1,\omega_2,\omega_3) = -D_2(\omega_1 + \omega_2 + \omega_3)^2 + j[2\,D_1(\omega_1 + \omega_2 + \omega_3) - 2D_3(\omega_1 + \omega_2 + \omega_3)^3] \qquad \textbf{(3.36), B2.19}$$

The required forward Volterra (current to photon) transfer functions, H_1 to H_3, can be derived from equations B2.17 – B2.19, following the same method as described in AppB1 with equations B1.24, B1.28 and B1.29. The expressions for the fundamental, IMD_2 and IMD_3 are the same as those in equations B1.32 to B1.34.

Appendix C: Matlab Simulation Results

(a) Precorrective case.



(b) Final corrected output

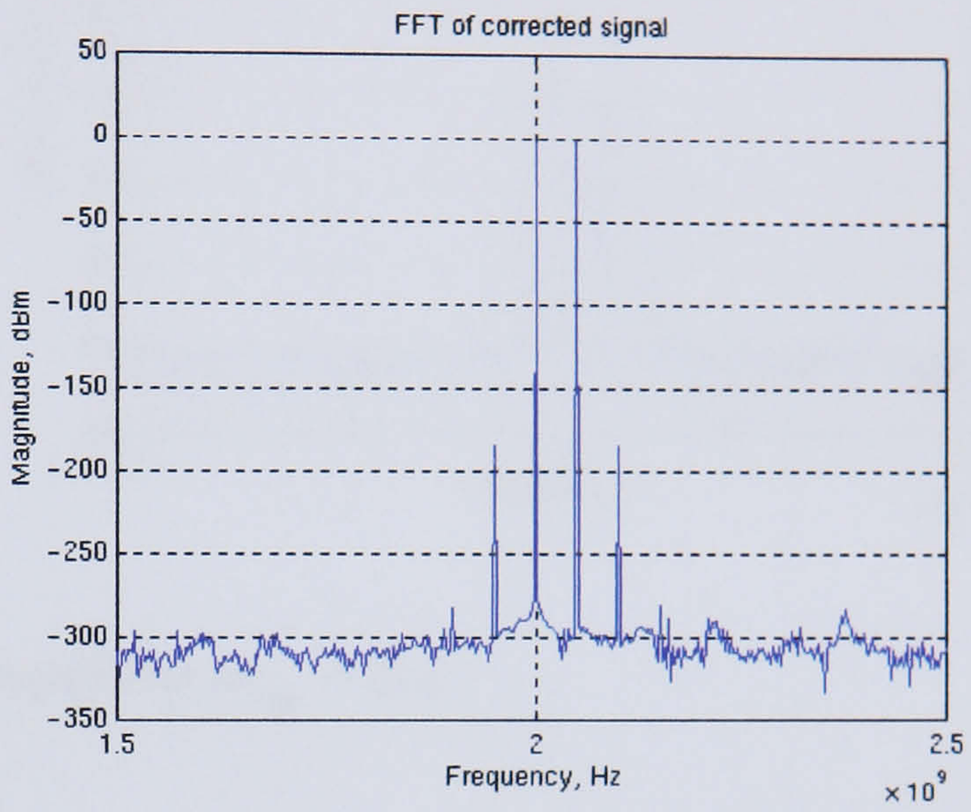
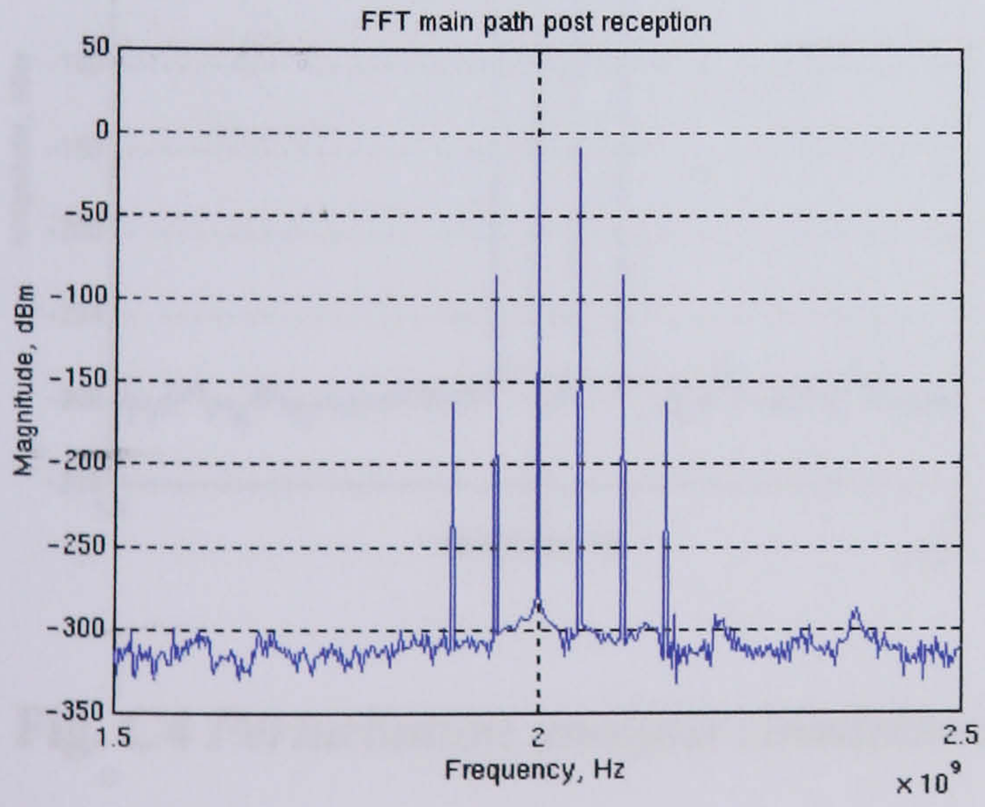


Fig. C1 *Taylor's series simulation results for* $m_{cur} = 0.2$

(a) Precorrective case.



(b) Final corrected output

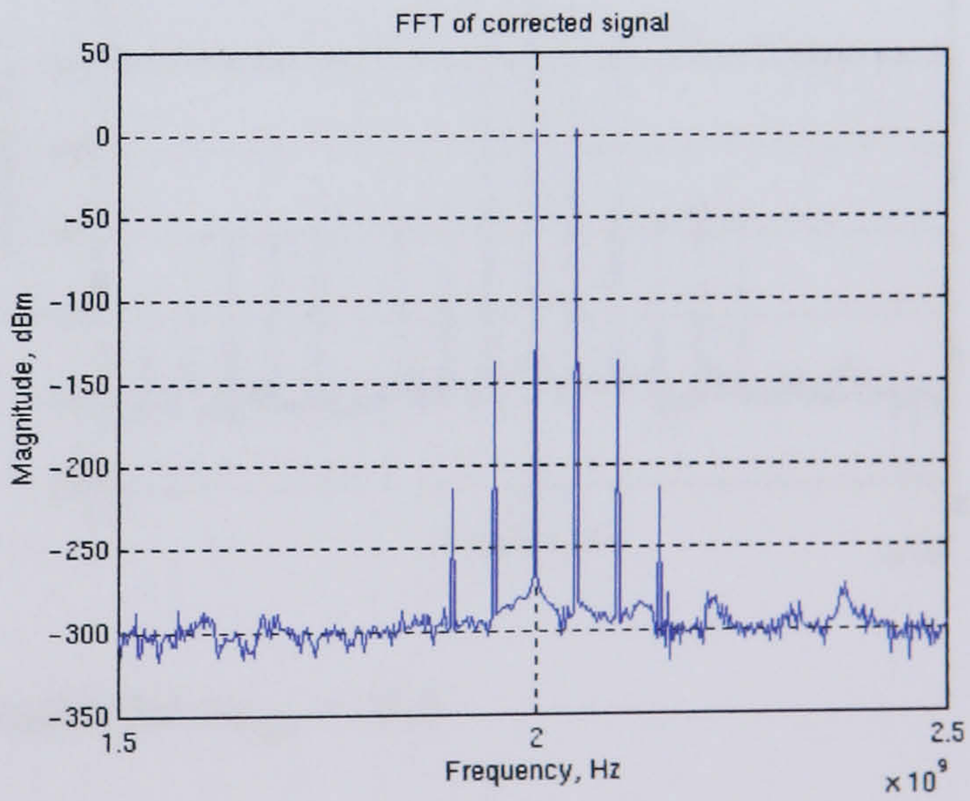


Fig. C2 *Taylor's series simulation results for* $m_{cur} = 0.6$

(a) Precorrective case.

(b) Final corrected output

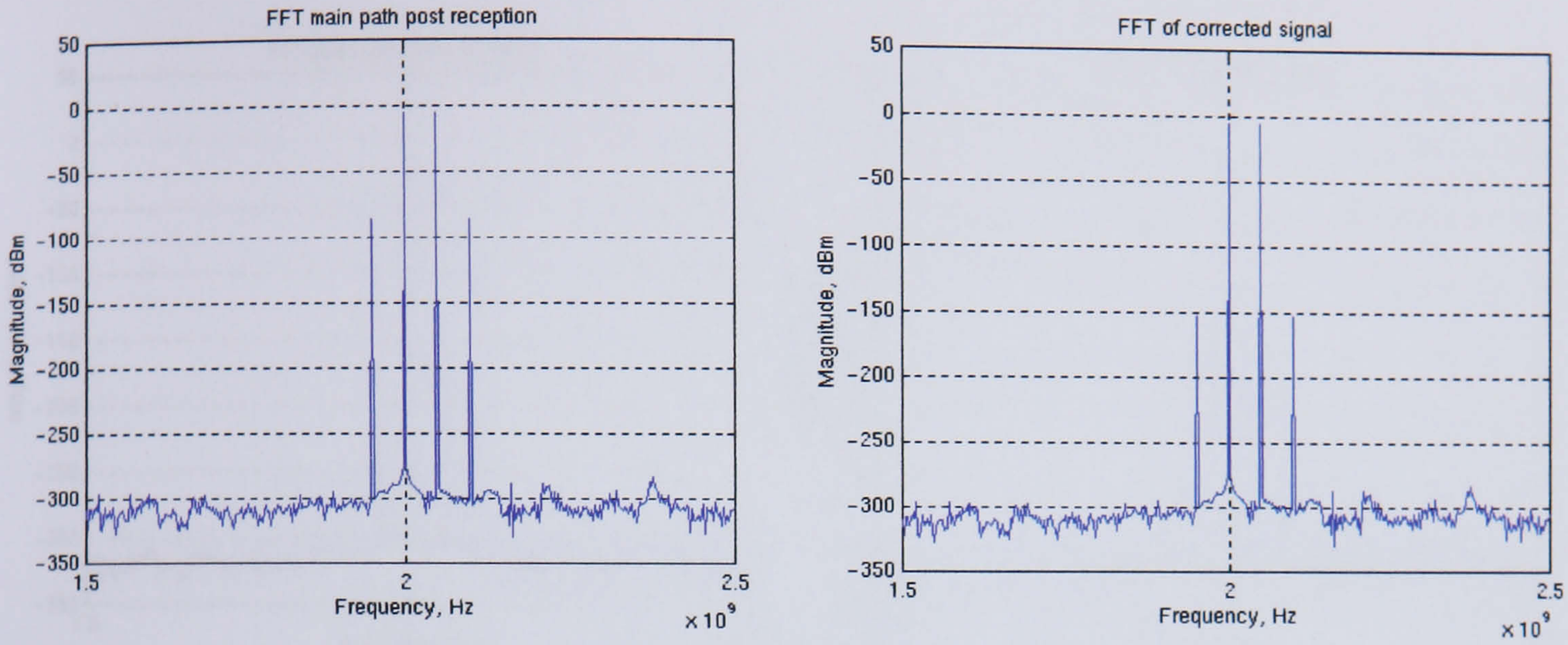


Fig. C3 *Perturbation analysis simulation results for* $m_{cur} = 0.2$

(a) Precorrective case.

(b) Final corrected output

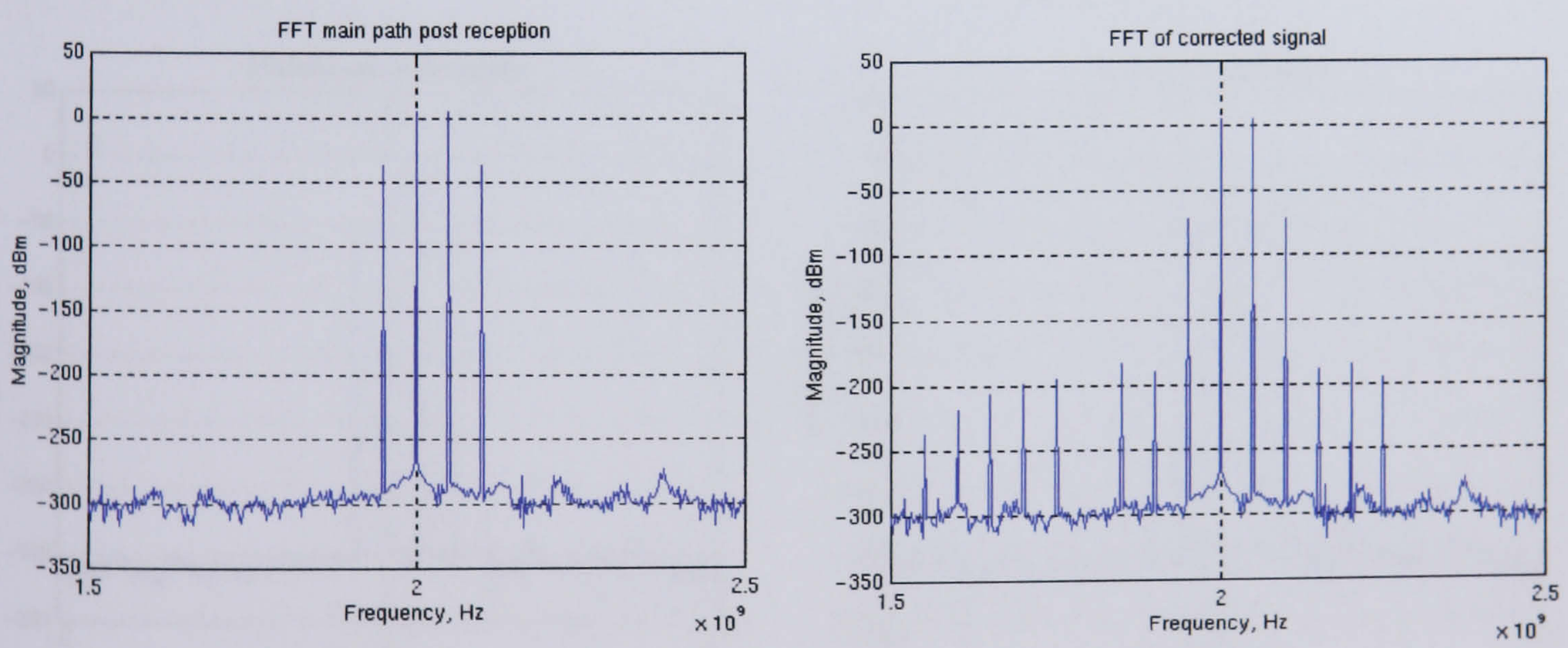


Fig. C4 *Perturbation analysis simulation results for* $m_{cur} = 0.6$

(a) Precorrective case.

(b) Final corrected output

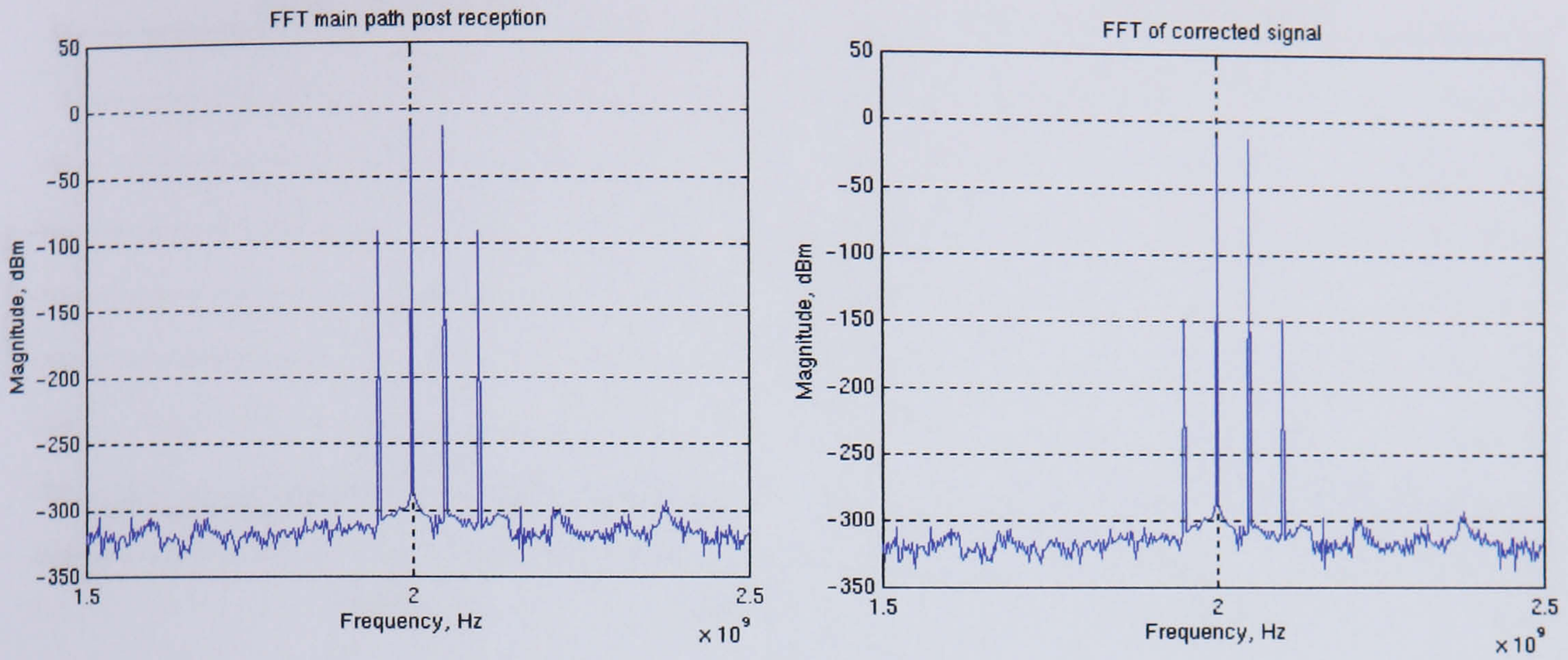


Fig. C5 *Volterra series analysis simulation results for* $m_{cur} = 0.2$

(a) Precorrective case.

(b) Final corrected output

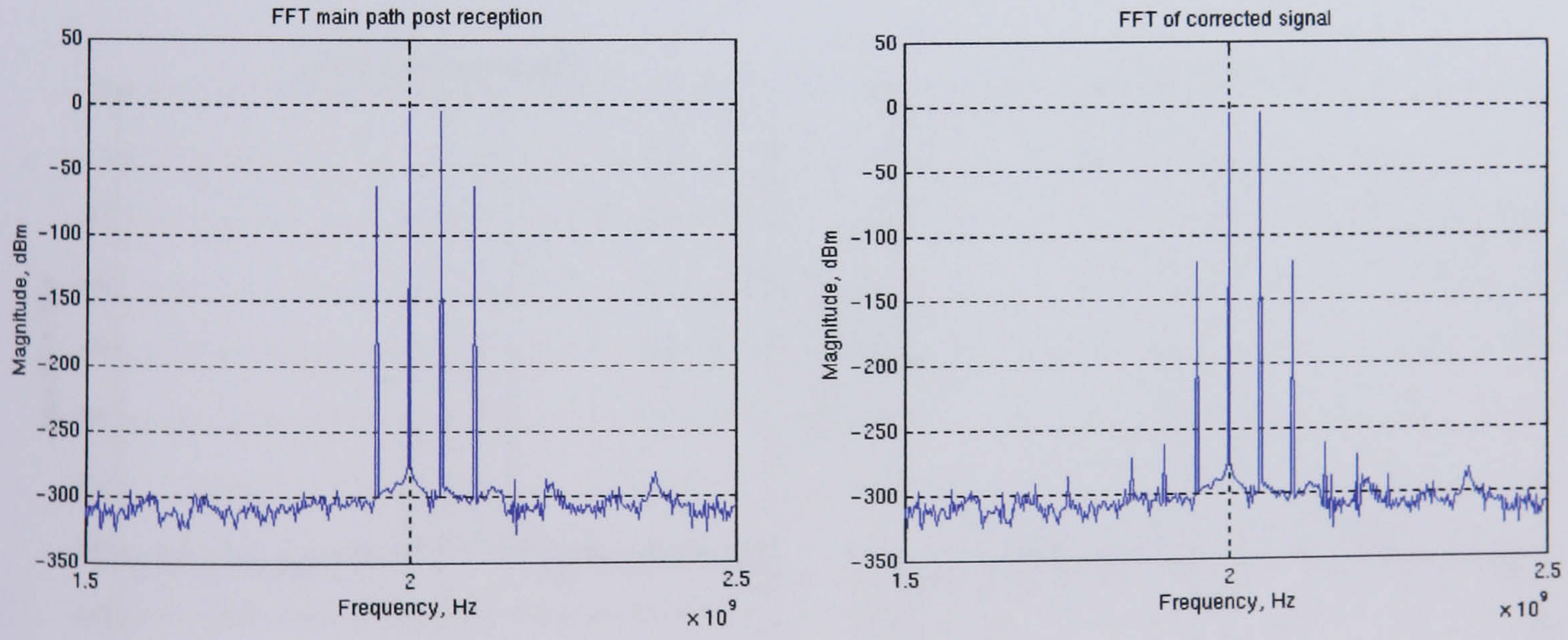


Fig. C6 *Volterra series analysis simulation results for* $m_{cur} = 0.6$

(a) Precorrective case.

(b) Final corrected output

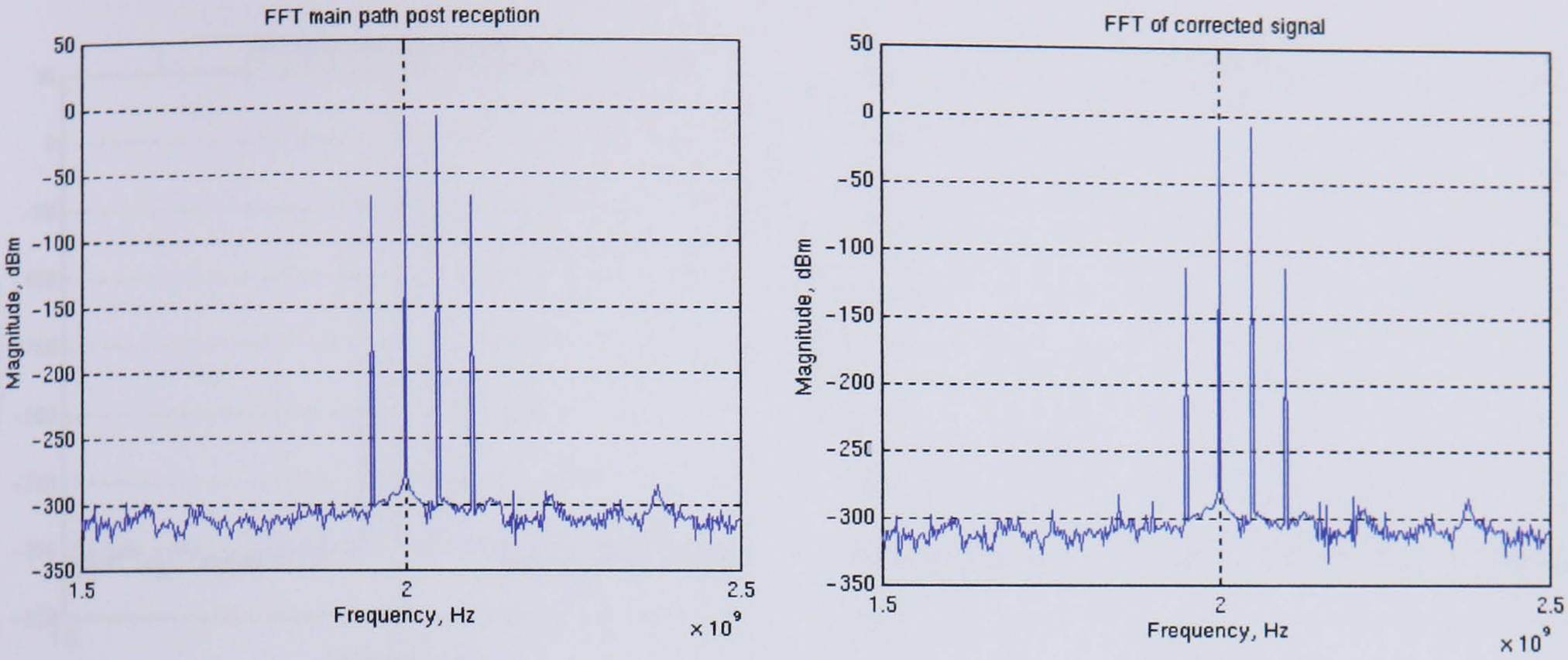


Fig. C7 *Advanced Volterra series analysis simulation results for* $R_{rd} = 1$, $T_{cap} = 60ps$ *and* $m_{cur} = 0.4$

(a) Precorrective case.

(b) Final corrected output

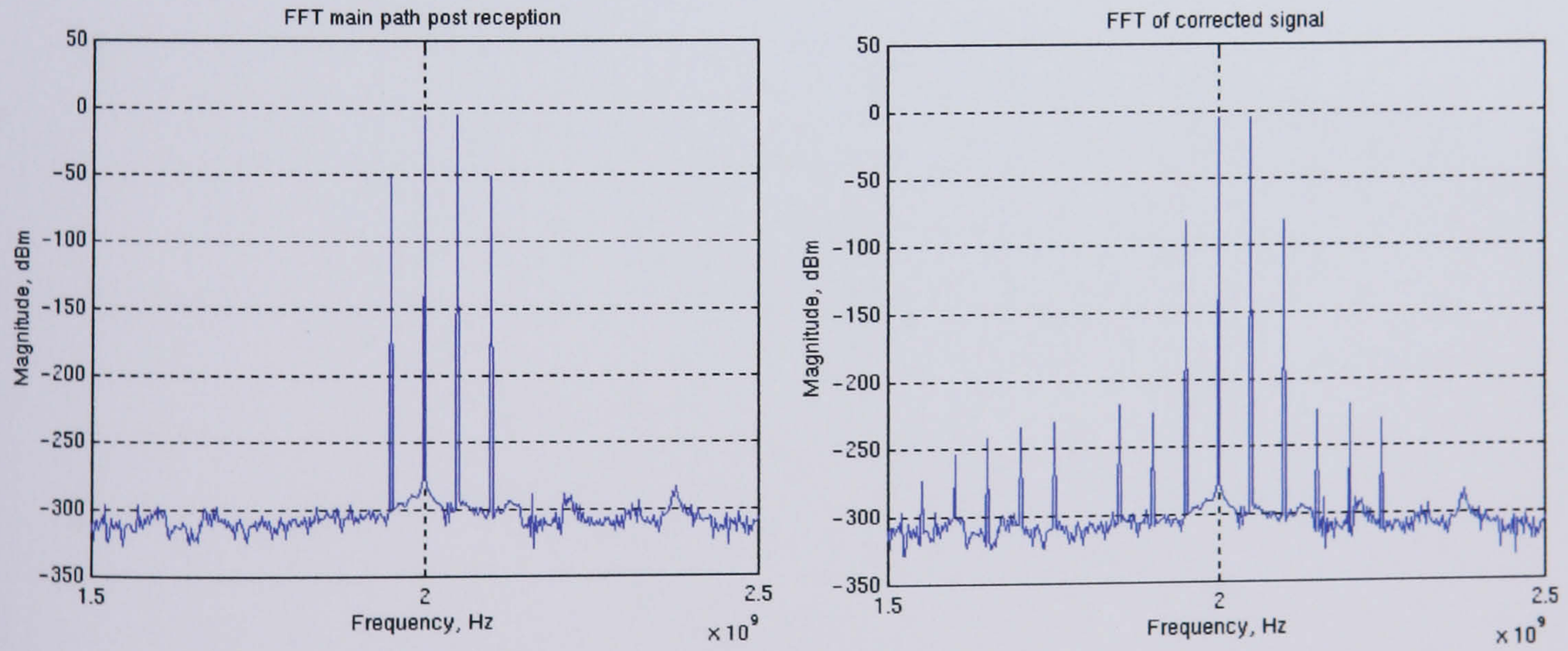


Fig. C8 *Advanced Volterra series analysis simulation results for* $R_{rd} = 5$, $T_{cap} = 1ps$ *and* $m_{cur} = 0.4$

(a) Precorrective case.

(b) Final corrected output

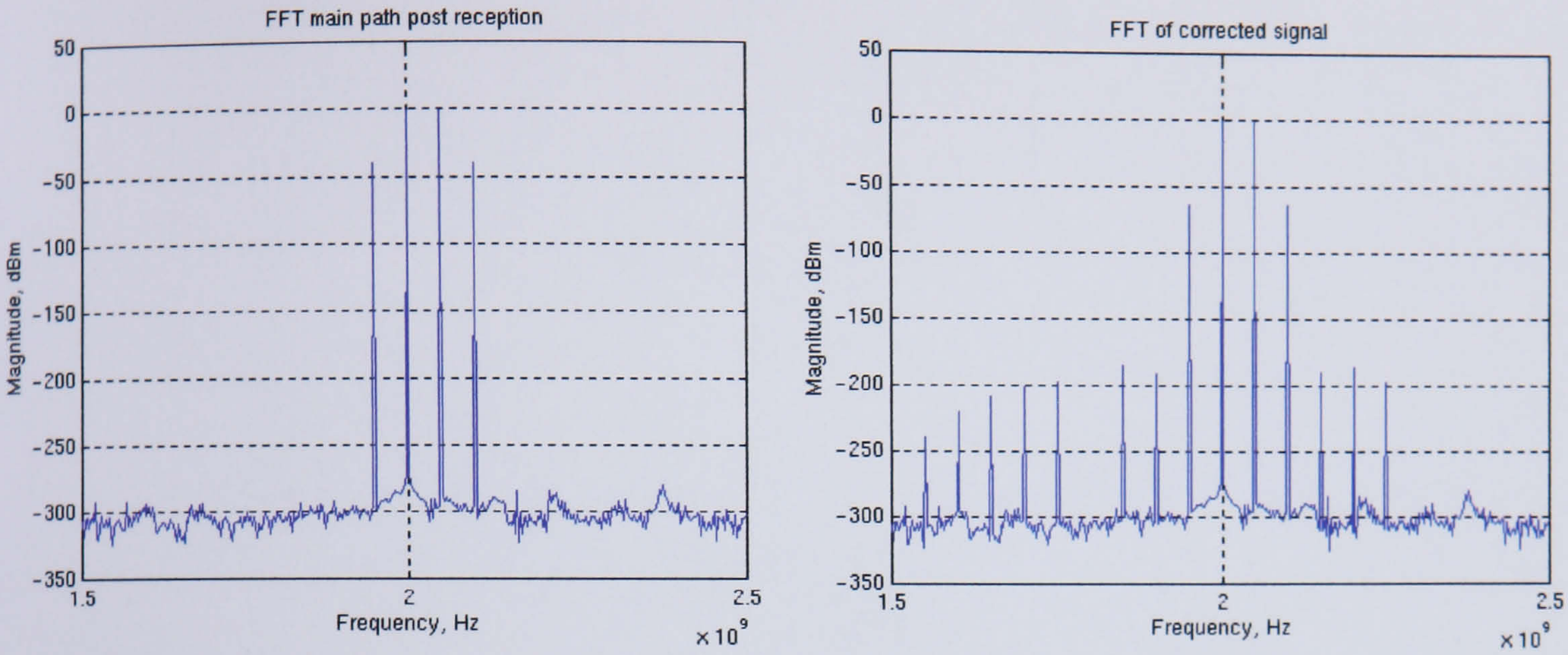
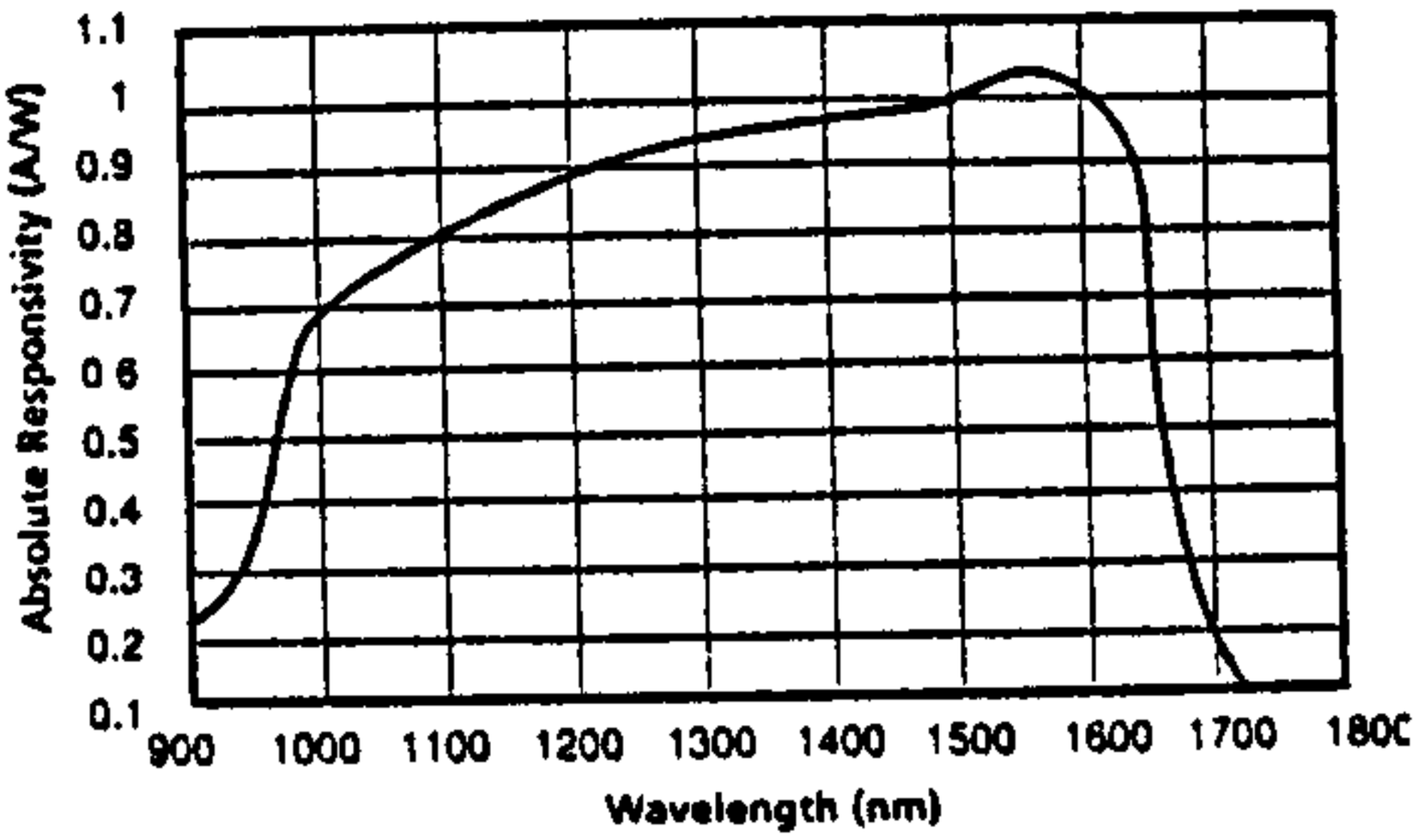


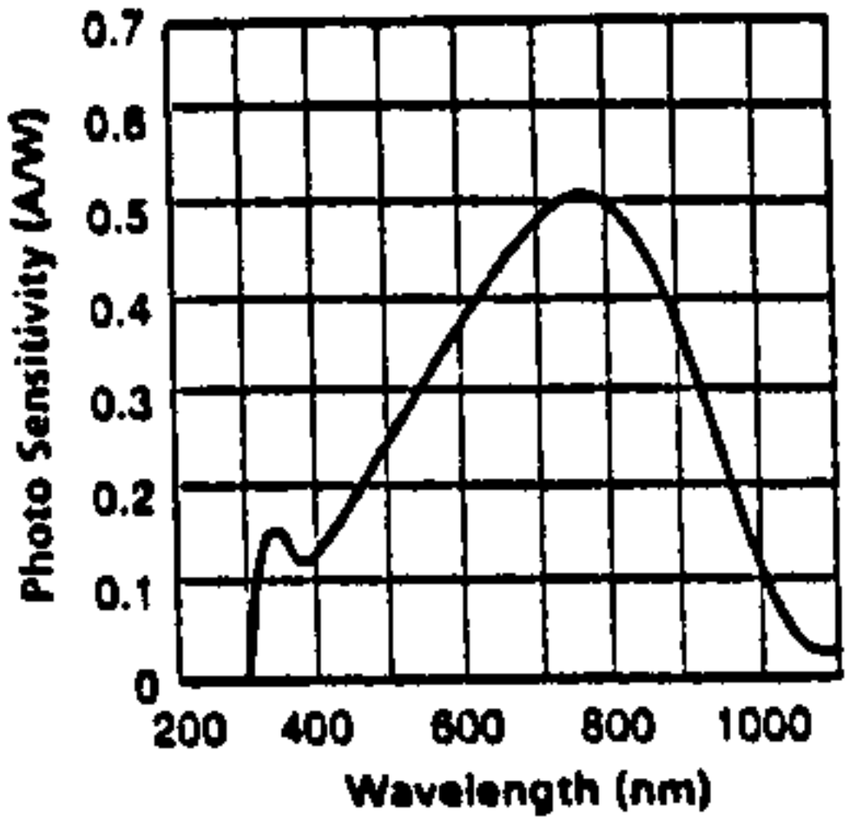
Fig. C9 *Advanced Volterra series analysis simulation results for* $R_{rd} = 5$, $T_{cap} = 1ps$ *and* $m_{cur} = 0.6$

Appendix D: Data sheets and circuit diagrams

Spectral Responsivity - ET-3000A



Spectral Response - E-2030A



Amplified Photodetector	ET-2030A
Detector Material	Silicon
Rise time	<500psec
Fall time	<500psec
Sensitivity @ 830nm	500 V/W
Sensitivity @ 1.3µm	N/A
Power Supply	240V AC
Frequency Response	75kHz - 1.2GHz
Active Area	0.12mm²
Acceptance Angle	10°
Equivalent Input Noise Current	7pA√Hz
Maximum Undistorted Output Voltage	500mVp-p
Connector	BNC
Mounting (tapped hole)	8/32 or M4

ABSOLUTE MAXIMUM RATINGS

Power-Supply Voltage (V_{CC} to V_{EE}) 12V
Input Voltage (IN_+, IN_-) ..... (V_{CC} + 0.3V) to (V_{EE} - 0.3V)
IN_ _ Current (Note 1) ..... ±10mA
Short-Circuit Duration (V_{OUT} to GND)
  V_{IN} < 1.5V ..... Continuous
  V_{IN} > 1.5V 0sec
Continuous Power Dissipation (T_A = +70°C)
  8-Pin SO (derate 5.88mW/°C above +70°C) 471mW
  8-Pin µMAX (derate 4.10mW/°C above +70°C) 330mW
  14-Pin SO (derate 8.33mW/°C above +70°C) 667mW
  16-Pin QSOP (derate 9.52mW/°C above +70°C) 762mW
Operating Temperature Range
  MAX41_ _E_ _ ..... -40°C to +85°C
Storage Temperature Range ..... -65°C to +160°C
Lead Temperature (soldering, 10sec) ..... +300°C

Stresses beyond those listed under "Absolute Maximum Ratings" may cause permanent damage to the device. These are stress ratings only, and functional operation of the device at these or any other conditions beyond those indicated in the operational sections of the specifications is not implied. Exposure to absolute maximum rating conditions for extended periods may affect device reliability.

ELECTRICAL CHARACTERISTICS

(V_{CC} = +5V, V_{EE} = -5V, T_A = T_{MIN} to T_{MAX}, unless otherwise noted. Typical values are at T_A = +25°C.) (Note 1)

PARAMETER	SYMBOL	CONDITIONS		MIN	TYP	MAX	UNITS
DC SPECIFICATIONS (R_L = ∞, unless otherwise noted)							
Input Offset Voltage	V_{OS}	V_{OUT} = 0V			1	8	mV
Input Offset Voltage Drift	TCV_{OS}	V_{OUT} = 0V			10		µV/°C
Positive Input Bias Current	I_{B+}	V_{OUT} = 0V, V_{IN} = -V_{OS}			3.5	20	µA
Negative Input Bias Current	I_{B-}	V_{OUT} = 0V, V_{IN} = -V_{OS}			3.5	20	µA
Input Resistance		IN+			500		kΩ
		IN-			30		Ω
Input Voltage Noise	e_n	f = 10kHz			2.2		nV/√Hz
Integrated Voltage Noise	E_{nRMS}	f = 1MHz to 100MHz			27		μV_{RMS}
Positive Input Current Noise	i_{n+}	f = 10kHz	MAX4112/MAX4117/ MAX4119		13		pA/√Hz
			MAX4113/MAX4118/ MAX4120		9		
Negative Input Current Noise	i_{n-}	f = 10kHz			14		pA/√Hz
Common-Mode Input Voltage	V_{CM}			-2.5		2.5	V
Common-Mode Rejection	CMR	V_{CM} = ±2.5V		45	50		dB
Power-Supply Rejection	PSR	V_S = ±4.5V to ±5.5V		60	80		dB
Open-Loop Transimpedance	Z_{OL}	V_{OUT} = ±2.0V, V_{CM} = 0V, R_L = 100Ω		250	500		kΩ
Quiescent Supply Current per Amplifier	I_{SY}	V_{IN} = 0V			5	6.5	mA
Output Voltage Swing	V_{OUT}	R_L = ∞		±3.5	±3.8		V
		R_L = 100Ω		±3.1	±3.5		
Output Current Drive	I_{OUT}	R_L = 30Ω, T_A = 0°C to +85°C		65	80		mA
AC SPECIFICATIONS (R_L = 100Ω, unless otherwise noted)							
Small Signal -3dB Bandwidth	BW_{SS}	V_{OUT} ≤ 0.1V_{RMS}	MAX4112/MAX4117		400		MHz
			MAX4113/MAX4119		270		
			MAX4118/MAX4120		300		

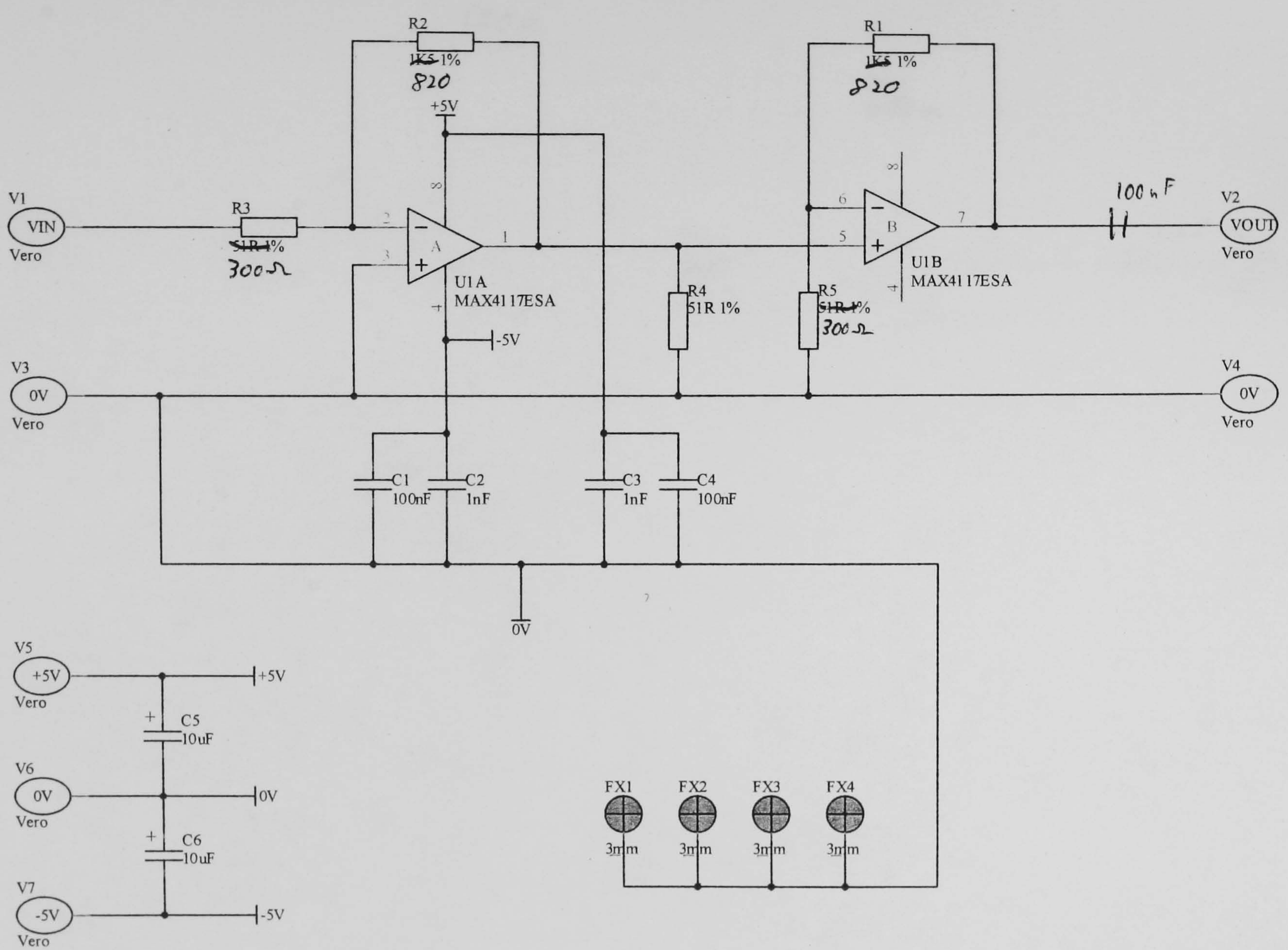
ELECTRICAL CHARACTERISTICS (continued)

(V_{CC} = +5V, V_{EE} = -5V, T_A = T_{MIN} to T_{MAX}, unless otherwise noted. Typical values are at T_A = +25°C.) (Note 1)

PARAMETER	SYMBOL	CONDITIONS		MIN	TYP	MAX	UNITS
AC SPECIFICATIONS (R_L = 100Ω, unless otherwise noted) (continued)							
0.1dB Gain Flatness	$BW_{0.1dB}$	MAX4112/MAX4117/MAX4119, A_{VCL} = +2			100		MHz
		MAX4113/MAX4118/MAX4120, A_{VCL} = +8			115		
Large-Signal -3dB Bandwidth	BW_{LS}	V_{OUT} = 2Vp-p	MAX4112/MAX4117		280		MHz
			MAX4119		145		
			MAX4113/MAX4118/MAX4120		240		
Slew Rate	SR	-2V ≤ V_{OUT} ≤ 2V	MAX4112/MAX4117/MAX4119		1200		V/µs
			MAX4113/MAX4118/MAX4120		1800		
Settling Time	t_S	to 0.1%, -1V ≤ V_{OUT} ≤ 1V	MAX4112/MAX4117/MAX4119		15		ns
			MAX4113/MAX4118/MAX4120		10		
		to 0.01%, -1V ≤ V_{OUT} ≤ 1V	MAX4112/MAX4117/MAX4119		35		
			MAX4113/MAX4118/MAX4120		25		
Rise/Fall Times	t_R, t_F	10% to 90%, -2V ≤ V_{OUT} ≤ 2V			3		ns
		10% to 90%, -50mV ≤ V_{OUT} ≤ 50mV			0.8		
Differential Gain	DG	f = 3.58MHz, R_L = 150Ω	MAX4112/MAX4117/MAX4119, A_{VCL} = +2		0.02		%
			MAX4113/MAX4118/MAX4120, A_{VCL} = +8		0.02		
Differential Phase	DP	f = 3.58MHz, R_L = 150Ω	MAX4112/MAX4117/MAX4119, A_{VCL} = +2		0.03		degrees
			MAX4113/MAX4118/MAX4120, A_{VCL} = +8		0.04		
Input Capacitance	C_{IN}				2		pF
Output Impedance	Z_{OUT}	f = 10MHz, A_{VCL} = +2			0.9		Ω
Spurious-Free Dynamic Range	SFDR	f_C = 5MHz, V_{OUT} = 2Vp-p	MAX4112/MAX4117/MAX4119, A_{VCL} = +2		-68		dBc
			MAX4113/MAX4118/MAX4120, A_{VCL} = +8		-62		
Two-Tone Third-Order Intercept	IP3	MAX4112/MAX4117/MAX4119, f_C = 10MHz, f_{C1} = 10.1MHz, A_{VCL} = +2			36		dB
Crosstalk		All hostile, V_{IN} = 1Vp-p, f = 10MHz			-75		dB

Note 1: The MAX4112/MAX4113/MAX4117–MAX4120 are designed to operate in a closed-loop configuration in which the IN- pin is driven by the OUT pin through an external feedback network. If an external voltage source is connected to IN-, current into or out of IN- must be limited to ±10mA, to prevent damage to the part.

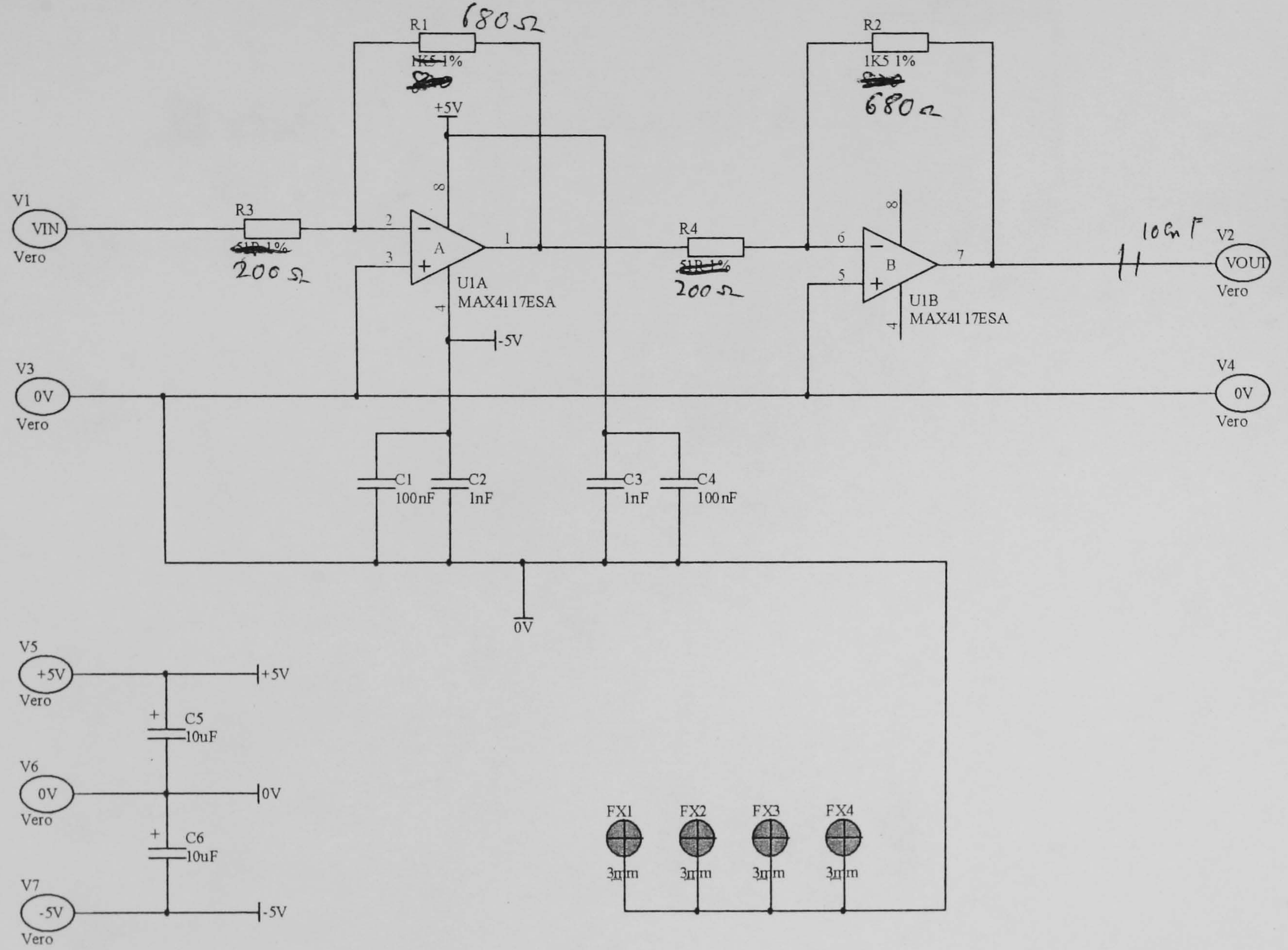
INV



Title: ***Current Feedback Amplifier (2)***			*University of Warwick* *School of Engineering* *Gibbet Hill Road* *Coventry* *CV4 7AL*
Size: A4	Number: 1383IG B	Revision: ISSUE 1	
Date: 30/01/2003	Time: 12:09:32	Sheet 1 of 1	
File: D:\PROTEL DXP\University\1383IG\1383IGB1.SchDoc			

THE UNIVERSITY OF WARWICK

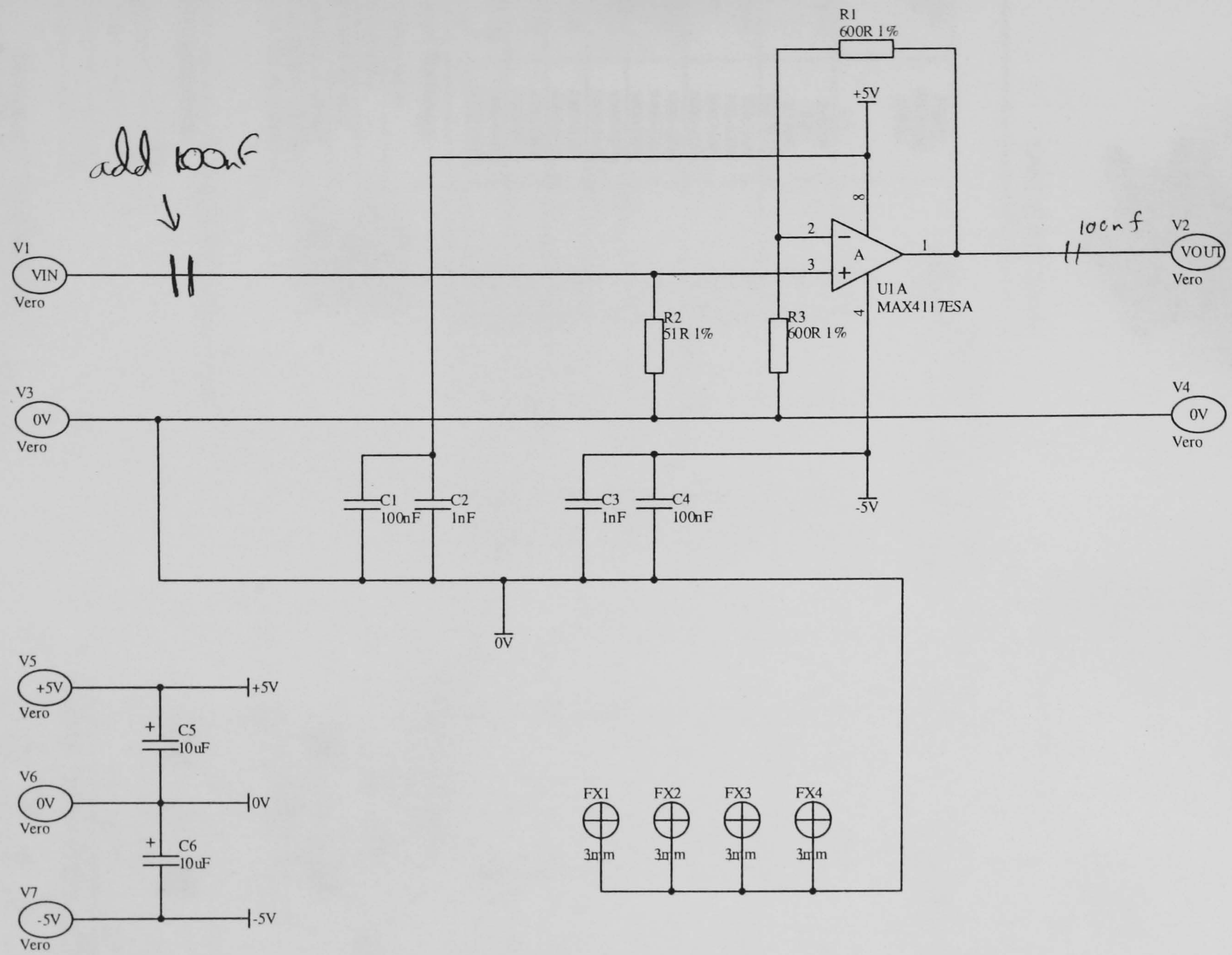
Non-Inv



Title: ***Current Feedback Amplifier***			*University of Warwick* *School of Engineering* *Gibbet Hill Road* *Coventry* *CV4 7AL*
Size: A4	Number: 1383IG A	Revision: ISSUE 1	
Date: 30/01/2003	Time: 12:08:51	Sheet 1 of 1	
File: D:\PROTEL DXP\University\1383IG\1383IGA1.SchDoc			

THE UNIVERSITY OF WARWICK

Buffer



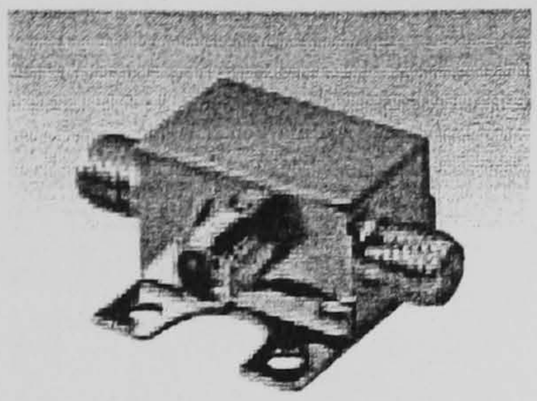
Title *Current Feedback Amplifier (3)*			*University of Warwick School of Engineering Gibbet Hill Road Coventry CV4 7AL*
Size: A4	Number: 1390IG A	Revision: ISSUE 1	
Date: 20/03/2003	Time: 12:57:49	Sheet 1 of 1	
File: D:\PROTEL DXP\University\1390IG\1390IGA1.SchDoc			

THE UNIVERSITY OF WARWICK

Coaxial

Power Splitters/Combiners ZX10-SERIES

2 Way-0°, 50Ω 2 to 12600 MHz



CASE STYLE: FL905

Features

- low insertion loss, 0.25 dB typ.
- excellent amplitude unbalance
- very good phase unbalance
- small size
- low cost
- patent pending

Applications

- communications
- defense
- cable tv relay
- radio
- cellular/GSM
- PCN/PCS
- VSAT
- DECT
- PHS
- UHF/VHF receivers/transmitters

Electrical Specifications (T_{AMB}=25°C)

MODEL NO.	FREQ. RANGE (MHz) f_L-f_U	ISOLATION (dB) Typ.	Min.	INSERTION LOSS (dB) above 3.0 dB Typ.	Max.	PHASE UNBALANCE (Deg.) Max.	AMPLITUDE UNBALANCE (dB) Max.	CASE STYLE	CONNECTION	PRICE $ Qty. (1-24)
ZX10-2-12	**2-1200**	**21**	**16**	**0.5**	**1.2**	**3.0**	**0.5**	FL905	ar	24.95
	2-20	29	21	0.3	0.7	3.0	0.5			
	20-600	25	18	0.3	0.8	2.0	0.4			
	600-1000	21	18	0.5	1.0	2.0	0.5			
ZX10-2-20	**200-2000**	**20**	**16**	**0.8**	**2.2**	**6.0**	**0.4**	FL905	ar	24.95
	800-1000	22	18	0.5	0.9	2.0	0.3			
	500-1500	22	18	0.5	1.3	3.0	0.4			
	1800-2000	20	18	1.6	2.2	6.0	0.4			
ZX10-2-25	**1000-2500**	**20**	**14**	**1.2**	**2.0**	**10.0**	**1.2**	FL905	ar	26.95
	1400-1800	18	16	0.9	1.2	6.0	0.6			
	1800-2000	19	16	0.9	1.2	6.0	0.6			
ZX10-2-42	**1900-4200**	**23**	**10**	**0.2**	**1.0**	**3.0**	**0.3**	FL905	ar	34.95
	2600-3400	23	17	0.2	0.5	2.0	0.3			
ZX10-2-71	**2950-7100**	**23**	**10**	**0.25**	**0.9**	**3.0**	**0.4**	FL905	ar	34.95
	4500-5700	23	18	0.2	0.4	3.0	0.3			
ZX10-2-98	**4750-9800**	**23**	**10**	**0.3**	**1.2**	**9.0**	**0.5**	FL905	ar	39.95
	7000-9000	23	18	0.3	0.8	8.0	0.4			
ZX10-2-126	**7400-12600**	**23**	**10**	**0.3**	**1.3**	**10.0**	**0.5**	FL905	ar	39.95
	9000-11000	23	16	0.3	0.6	5.0	0.3			

Maximum Ratings

Operating Temperature	-40°C to 85°C
Storage Temperature	-55°C to 100°C
Power Input (as a splitter)	1.0*W max.
Power Input (as a combiner)	0.125**W max.

*0.5W max. for ZX-2-10, ZX-2-20
**0.1W max. for ZX-2-42 & ZX-2-126

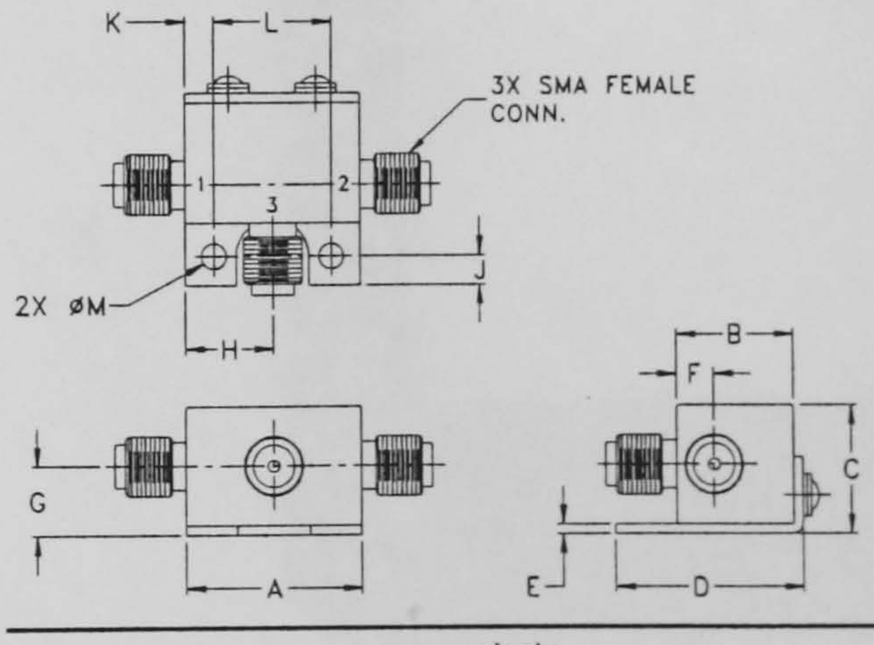
Port Connections

PORT	(ar)
SUM PORT	3
PORT 1	1
PORT 2	2

designers kit available

KIT No.	No. of Units in KIT	Description	Price $ per KIT
K1-ZX10	7	1 of each model in free deluxe wood storage case	199.95

Outline Drawing



Outline Dimensions (inch / mm)

A	B	C	D	E	F	G	H	J
.74	.50	.54	.80	.04	.16	.29	.37	.122
18.80	12.70	13.72	20.32	1.02	4.06	7.37	9.40	3.10

K	L	M	wt.
.122	.496	.106	grams
3.10	12.60	2.69	20.0

REV. A
M84793
ZX10-2-12 ED-10145/1
ZX10-2-20 ED-10145/2
ZX10-2-25 ED-10145/3
ZX10-2-42 ED-10281/2
ZX10-2-71 ED-10146/1
ZX10-2-98 ED-10280/1
ZX10-2-126 ED-10227/1
ZX10-SERIES
HY/RS/CP
021230

Mini-Circuits®
INTERNET http://www.minicircuits.com
P.O. Box 350166, Brooklyn, New York 11235-0003 (718) 934-4500 Fax (718) 332-4661
Distribution Centers NORTH AMERICA 800-654-7949 • 417-335-5935 • Fax 417-335-5945 • EUROPE 44-1252-832600 • Fax 44-1252-837010